

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

VOLUME 58, ART. 6. PAGES 721-970

Editor

ROY WALDO MINER

PROPERTIES OF SURFACES

BY

CECIL V. KING (*Conference Chairman*), R. J. ARCHER, J. A. BECKER, W. H. BRATTAIN, C. J. CALBICK, B. H. CARROLL, E. N. CLARKE, D. G. FLOM, E. E. FRANCOIS, C. G. B. GARRETT, E. A. GULBRANSEN, J. H. HONIG, V. K. LAMER, J. T. LAW, H. H. PODGURSKI, T. N. RHODIN, JR., R. H. SAVAGE, H. A. TAYLOR, H. H. UHLIG, W. WEST, AND D. L. WHITCOMB.

Consulting Editor

CECIL V. KING



NEW YORK
PUBLISHED BY THE ACADEMY
September 15, 1954

THE NEW YORK ACADEMY OF SCIENCES

(Founded in 1817)

COUNCIL, 1954

President

JOHN TEE-VAN

President-Elect

MAURICE L. TAINTER

Vice-President

ELMORE H. NORTHEY

Recording Secretary

ROSS F. NIGRELLI

Corresponding Secretary

JUNIUS BIRD

Treasurer

DONALD M. BENJAMIN

Editor

ROY WALDO MINER

Elected Councilors

1952-1954

WILLIAM H. COLE

CHARLES H. TOWNES FREDERICK C. NACHOD

1953-1955

EDWARD J. KEMPF
BORIS PREGEL

CHARLES D. MARPLE
JOHN TURKEVICH

1954-1956

JOHN M. CONVERSE
RANDOLPH T. MAJOR

B. M. DUGGAR
ABRAHAM SLAVIN

Finance Committee

HARDEN F. TAYLOR, *Chairman*

GORDON Y. BILLARD

ROBERT F. LIGHT

Executive Director

EUNICE THOMAS MINER

SECTION OF GEOLOGY AND MINERALOGY

THOMAS N. WALTHIER, *Chairman*

ANGELINA ROSE MESSINA, *Secretary*

SECTION OF BIOLOGY

M. J. KOPAC, *Chairman*

HILARY KOPROWSKI, *Secretary*

DIVISION OF MYCOLOGY

SAMUEL M. PECK, *Chairman*

SECTION OF PSYCHOLOGY

NEWMAN L. HOOPINGARNER, *Chairman*

ALBERTA S. GILINSKY, *Secretary*

SECTION OF ANTHROPOLOGY

WILLIAM L. THOMAS, JR., *Chairman*

JOHN L. LANDGRAF, *Secretary*

SECTION OF PHYSICS AND CHEMISTRY

CECIL V. KING, *Chairman*

J. J. DENTON, *Secretary*

SECTION OF OCEANOGRAPHY AND METEOROLOGY

ERNEST J. CHRISTIE, *Chairman*

MAYNARD E. SMITH, *Secretary*

SECTION OF MATHEMATICS AND ENGINEERING

SEBASTIAN B. LITTAUER, *Chairman*

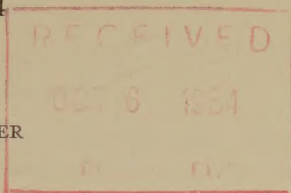
DAVID B. HERTZ, *Secretary*

These Sections and the Division hold meetings regularly, one evening each month, during the academic year, October to May, inclusive. Two-day conferences also are held at irregular intervals. All meetings are held at the building of The New York Academy of Sciences, 2 East Sixty-third Street, New York 21, New York.

September 15, 1954

Editor

ROY WALDO MINER



PROPERTIES OF SURFACES*

Conference Chairman and Consulting Editor

CECIL V. KING

CONTENTS

The Life History of Adsorbed Atoms, Ions, and Molecules. By JOSEPH A. BECKER . . .	723
Absorbent-Adsorbate Interactions and Surface Heterogeneity in Physical Adsorption. By J. M. HONIG	741
The Kinetics of Surface Properties. By H. AUSTIN TAYLOR	798
The Effect of Monolayers on the Rate of Evaporation of Water. By ROBERT J. ARCHER AND VICTOR K. LAMER	807
Classical Theory of Diffusion and the Oxidation of Metals. By EARL A. GULBRANSEN	830
Passivity and Adsorption. By HERBERT H. UHLIG	843
Oxide Film Composition Studies. By THOR N. RHODIN, JR.	855
Surface Studies with the Electron Microscope. By C. J. CALBICK	873
The Adsorption of Dyes to Microcrystals of Silver Halide. By W. WEST, B. H. CARROLL, AND D. L. WHITCOMB	893
Adsorption and Exchange in Metal-Metal Ion Systems. By CECIL V. KING	910
The Adsorption of Gases and Vapors on Germanium. By J. T. LAW AND E. E. FRANCOIS	925
Electrical Conductivity of Germanium Surfaces. By EDWARD N. CLARKE	937
Exploration of Metal Surfaces with Fine Wires. By R. H. SAVAGE AND D. G. FLOM . .	946
Surface Properties of Germanium and Silicon. By W. H. BRATTAIN AND C. G. B. GARRETT	951
Stabilization of Metal Carbides by Nonmetallic Elements. By HARRY W. PODGURSKI	959

* This series of papers is the result of a conference on *Properties of Surfaces* held by the Section of Physics and Chemistry of The New York Academy of Sciences, November 5 and 6, 1953.

Copyright, 1954, by The New York Academy of Sciences

THE LIFE HISTORY OF ADSORBED ATOMS, IONS, AND MOLECULES

By Joseph A. Becker

Bell Telephone Laboratories, Murray Hill, N. J.

Recent work with the field emission electron microscope has greatly increased our knowledge of some of the processes which take place when foreign atoms or molecules are adsorbed on metal surfaces.^{5, 12} In particular, this work has emphasized that the electron work function of a metal varies from one crystallographic plane to the next and that, in adsorption phenomena, the arrangement of the metal atoms on the surface plays an important part. This microscope, under certain special conditions, seems to reveal the location of individual oxygen atoms in some oxygen molecules adsorbed on a tungsten surface. Another new technique, the "flash filament" and ion gauge, makes it possible to determine small amounts of gaseous atoms adsorbed on a few cm.² of metal surface at pressures near 10^{-7} mm. Hg.⁶ As a result, it is possible to measure evaporation rates and sticking probabilities as a function of temperature and amount adsorbed. Because of these and other advances in our understanding of adsorption phenomena, it has seemed advisable to bring up to date *The Life History of Adsorbed Atoms and Ions* published in 1929.

In this revised story, we shall deal mostly with two rather simple adsorption systems, namely cesium on tungsten and oxygen on tungsten. Tungsten is chosen as the adsorbent surface because it is readily available in wire or ribbon form, can be purified by heating to very high temperatures, and because its properties are well understood. Cesium is chosen as the one adsorbate because: (1) it decreases the electron work function drastically; (2) its evaporation rate as ions and as atoms can be determined readily; (3) its vapor pressure can be easily varied so that a monolayer is deposited in from one second to 500 seconds; and (4) every cesium atom that strikes the surface sticks to it. Oxygen is chosen as the second case because it is one of the simpler molecules. It is strongly chemisorbed on tungsten but is only slightly absorbed; and it increases the electron work function drastically.

It is natural to ask what is meant by an adsorbed atom, ion, or molecule and how can they be distinguished. We shall abbreviate these terms to adatoms, adions, and admols. If the valence electron of the cesium is still associated with its nucleus when the cesium is adsorbed, we shall call it an adatom. This will still be so if the valence electron is shared with one or more tungsten atoms or if it and another electron from the tungsten form a paired electron bond. On the other hand, if the valence electron is attracted more strongly by tungsten atoms than it is by its own nucleus, and is transferred to the tungsten metal, then we shall call it an adion. Adatoms and adions can be distinguished by the electrical fields or potentials they produce and by the way in which these fields affect the emission of electrons, ions, and atoms. A positive adion should induce a negative charge on the tungsten atoms and act like a positive dipole. An array of such dipoles can produce fields of the

order of 100 million volts/cm. near the surface. These fields are in a direction to make it easier for electrons to escape from the surface, but to make it harder for positive adions to escape. Adatoms, particularly if they form paired electron bonds, would be expected to produce much smaller fields and consequently should affect the emission of electrons and ions only slightly.

An admol or adsorbed molecule is defined as an aggregate of two or more atoms adsorbed on a metal surface in such a way that the attractive forces between the atoms for each other is greater than the attractive forces of the atoms to the metal. The aggregate acts as a unit and if one atom in it is displaced it pulls the other atom or atoms with it. In the case of oxygen on tungsten, if the O—O bond strength on the surface exceeds the O₂—W bond strength, the oxygen will exist as admols; if two O—W bonds greatly exceed the O—O bond, then an admol is likely to split up into two adatoms and a considerable amount of heat will be liberated. Whether or not such splitting up occurs on adsorption can be determined with "tagged" molecules.¹ The tagging can consist of isotopic or of radioactive molecules. If no interchange of partners occurs, it is concluded that the atoms remained associated as a molecule throughout their time of adsorption and were admols continuously. If interchange of partners occurs, then it is concluded that for all or part of the time adatoms existed. From such experiments it has been found:¹ (1) that for physical adsorption, in which Van der Waals or polarization forces are dominant, molecules become admols and the heat of adsorption corresponds to about 0.1 e.v. (electron volts equivalent); (2) that for chemisorption, molecules become adatoms by transferring electronic bonds and the heat of adsorption corresponds to 1.0 to 5.0 e.v. Recently, the field emission microscope has given pictures from which we deduce that *under certain circumstances* admols of oxygen exist on tungsten because they act as a unit.

Cesium Adsorbed on Tungsten

The properties of cesium and other alkali metals adsorbed on tungsten were actively investigated about 25 years ago. Langmuir and Kingdon showed that, when cesium atoms strike a hot tungsten surface, every atom is momentarily adsorbed and is evaporated as a positive ion.⁹ It is thus possible to measure the arrival rate rather precisely. If the tungsten temperature is low enough, every atom that strikes the surface sticks to it and becomes adsorbed. When the amount adsorbed reaches a high enough value, evaporation begins and increases until its rate equals the arrival rate. Langmuir also showed that the adsorbed cesium increased the thermionic electron emission enormously and hence deduced that the electron work function was drastically decreased. Ives⁸ and others showed that the photoelectric emission led to the same conclusion. These experiments showed that, as the amount of adsorbed cesium, θ , increases, the electron work junction, φ_e , decreases to a minimum value and then increases toward the value for bulk cesium. We will define $\theta = 1.0$ when φ_e is a minimum.

Soon thereafter Becker² devised a method for measuring θ over a considerable range of arrival rates and tungsten temperatures. It was found that

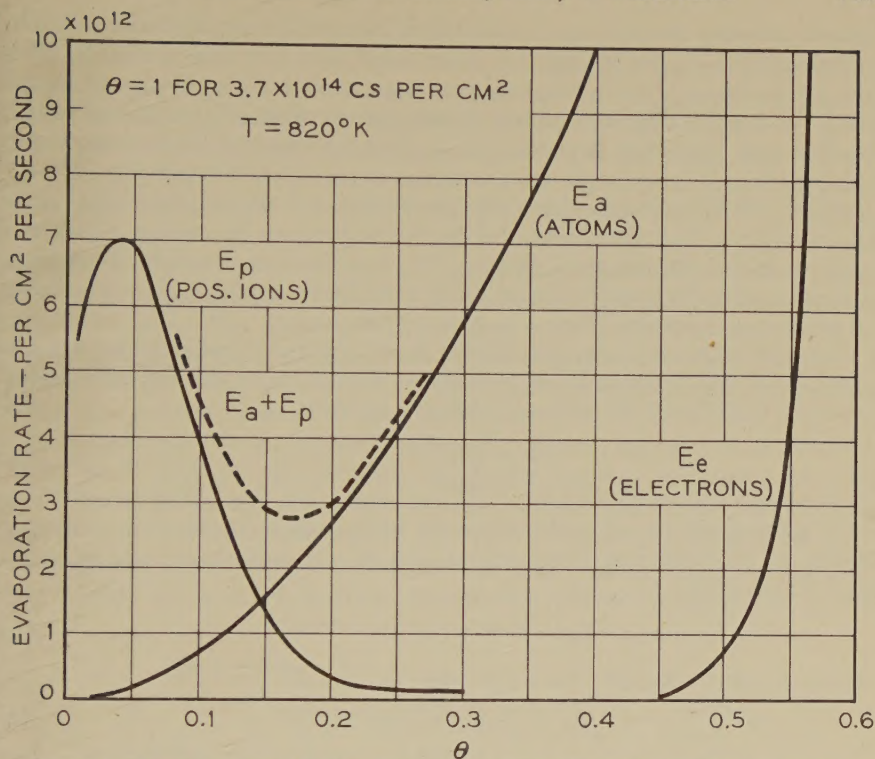


FIGURE 1. Evaporation rate of electrons, atoms, and ions versus amounts of Cs adsorbed: $T = 820^{\circ}\text{K}$.

$\theta = 1.0$ when about 3.7×10^{14} cesium atoms are adsorbed per cm^2 of tungsten surface. This is nearly equal to $\frac{1}{4}$ the number of tungsten atoms per cm^2 on a 110 plane. He confirmed Langmuir's conclusion that the sticking probability for Cs on W was unity. It was then possible to determine the rate of emission of electrons, atoms, and ions as a function of θ at constant temperature. FIGURE 1 shows a plot of these evaporation rates E versus θ at $T = 820^{\circ}\text{K}$. The subscripts p , a , and e refer to positive ions, atoms, and electrons, respectively. The electron emission increases very rapidly with θ ; the atom emission increases slowly with θ ; the positive ion emission increases with θ from 0 to about 0.04, comes to a maximum and then decreases as θ increases. From the shapes of the curves, one can conclude that φ_e decreases rapidly with θ , that φ_a changes only slowly with θ , but that φ_p increases as θ increases.

In the range of θ from 0.05 to 0.25, both atoms and ions evaporate at appreciable rates from the same surface at the same temperature. From this fact, we deduce that both atoms and ions are adsorbed on the surface. We reason that the forces tending to remove the cesium valence electron are likely to increase as the atom approaches the surface up to the distance of closest approach. At this distance, the transfer of electrons between tungsten and cesium nuclei may occur so rapidly that it would be meaningless to distinguish between

atoms and ions. However, we believe that the experimental facts are explained more readily or more simply if it is assumed that the time during which a cesium nucleus has no valence electron associated with it, is long compared with the time it takes an electron to traverse an orbit. A particular cesium particle may then exist as an adatom or adion for times which are comparable with the time it takes an electron or an ion to move a few angstroms from the surface. By assuming that both adatoms and adions exist, we can readily make quantitative calculations of the fields that exist near the surface which are attributable to the adsorbed cesium. We can also compute how these fields and the potentials associated with them will affect the probability of emission of electrons, atoms, and ions.

In order to convert rates of emission, E , to work functions, φ , we must know the relationship between these quantities. For electrons this can be obtained from the well-known Richardson-Dushman equation

$$i = 120T^2 e^{-\varphi_e e/kT} \quad (1a)$$

in which i = amperes/cm.²; T = °K; φ_e = electron work function in volts; e = electron charge in coulombs; k = Boltzmann's constant in joules per degree. Converting i to E_e or electron emission per cm.² per sec. and using the base 10 we obtain

$$E_e = 8 \times 10^{20} T^2 10^{-\varphi_e 5050/T} \quad (1b)$$

The analogous equations for atoms and for ions are

$$E_a = N_a 10^{13} T^{1/2} 10^{-\varphi_a 5050/T} \quad (2)$$

$$E_p = N_p 10^{13} T^{1/2} 10^{-\varphi_p 5050/T} \quad (3)$$

in which N_a and N_p are the number of adatoms and adions, respectively per cm.²; $10^{13} T^{1/2}$ is the number of collisions a surface particle makes per sec. with its neighbors; the exponential factors give the probability that a collision will give the particle enough energy to escape. From experimental values of E_e for various θ and T and EQUATION 1b we can calculate φ_e as a function of θ . The analogous calculations for φ_a and φ_p cannot be made until values of N_a and N_p can be determined. This can be done from a theoretical analysis of the φ_e versus θ curves.

In its simplest form this analysis assumes that the adsorbed cesium acts like a sheet of positive charge of σ coulombs/cm.² at a distance ℓ cm. above the tungsten surface. This induces a charge of $-\sigma$ at a distance of ℓ below the surface. The field inside this condenser is $300 \times 4 \pi \sigma$ volts/cm. The work an electron must do to escape from the tungsten is reduced by $1200 \pi \sigma \ell$. On the basis of the assumption stated above, the charge σ will be attributable almost entirely to positive adions; uncharged adatoms, particularly if they are attributable to paired electron bonds, will have only a small dipole moment compared to that for adions. Hence $\sigma = N_p e$ and

$$\Delta\varphi_e = -1200 \pi N_p e \ell = -1.8 \times 10^{-6} N_p \ell \quad (4)$$

The next step is to derive values for ℓ . To do this we must define more carefully what is meant by the "surface" and investigate how far above the surface the cesium nucleus will be. It soon is apparent that the answer will be different for different planes of the tungsten crystal. FIGURE 2 shows a top and a section view for a cesium ion resting on the 100 and 110 plane. The size and location of tungsten atoms and cesium ions are taken from crystallographic data and Pauling's book.¹³ A cesium ion will attract electrons in the tungsten. This extra surface charge will be found as far from the interior as electrons can go, which will be a distance of one radius from the tungsten nucleus. The electronic surface will be a plane which is tangent to the radii of the outermost tungsten atoms. The distance from this plane to the cesium nucleus is taken as ℓ . For the 100 plane $\ell = 0.8 \text{ \AA}$. while for the 110 plane it is $1.2 \text{ \AA}$. or about 1.5 times as large. Hence an ion on an 110 plane should be about 1.5 times as effective in reducing φ_e as an ion on the 100 plane.

Unfortunately, the φ_e versus θ curves for different planes are not known. However, it is known that for clean tungsten φ_e for the 110 plane is 5.5 volts or higher while for the 100 plane it is about 4.6 volts. From the work of Martin¹¹ for Cs on W it follows that near $\theta = 1.0$, φ_e for the 110 plane is lower than for any other plane and in particular lower than φ_e for the 100 plane. Hence, near the optimum activity, the emission will be dominated by the 110

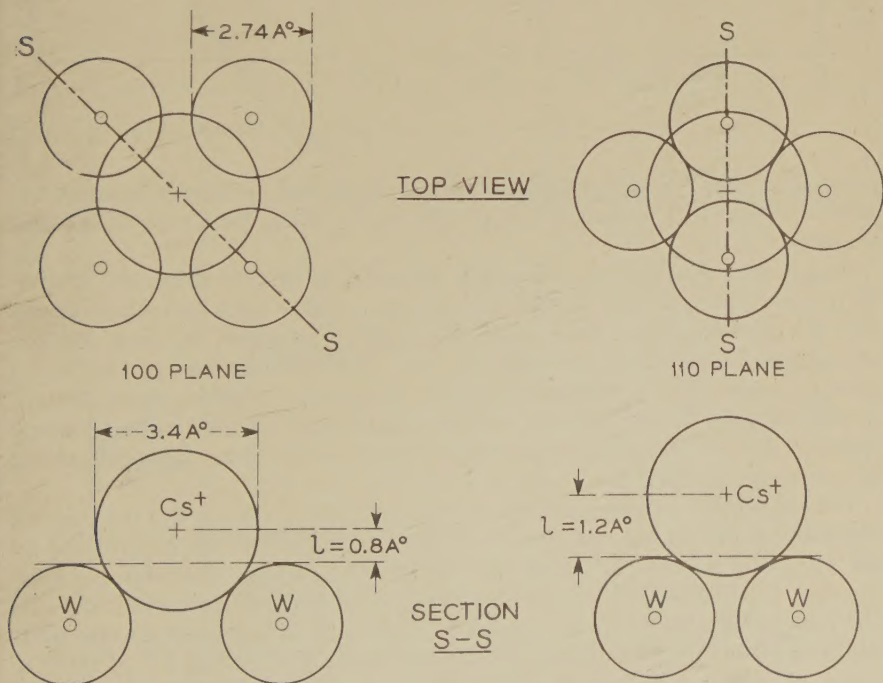


FIGURE 2. Location of cesium ions on the 100 and 110 planes of tungsten.

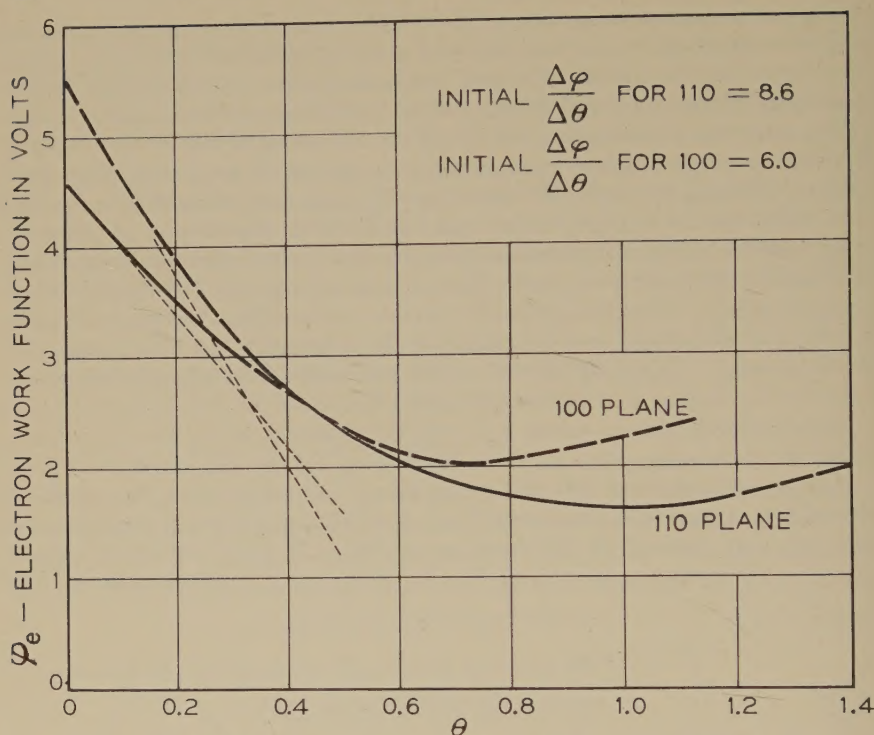


FIGURE 3. Electron work function versus amount adsorbed for 100 and 110 planes.

plane while near $\theta = 0$, the emission will be dominated by regions near the 100 plane. From other work it is known that the lowest ϕ_e for Cs on W is about 1.6 volts.

Hence, it is possible to draw fairly accurate curves for these two planes. Such curves are drawn in FIGURE 3. In them, regions for which experimental data are good are drawn solid, while other regions for which the data are poor are dashed. In sketching the curve for the 100 plane near its optimum θ , we have indicated that its optimum probably occurs at a smaller cesium concentration than for the 110 plane, which has been used to define $\theta = 1$. This reversal of the emission pattern between θ near zero and θ near the optimum has been frequently observed.

For our present purpose, the most significant part of FIGURE 3 is the value of the initial slopes for the two planes. It is to be noted that the initial slope for the 110 plane is about 1.5 times that for the 100 plane as was predicted from theory. Furthermore if we use the values of ℓ deduced from FIGURE 2 and the values of $\Delta\phi_e$ from FIGURE 3 and substitute them in EQUATION 4 to calculate N_p for say $\theta = 0.10$ we find for 110P, $N_p = 0.86 / (1.8 \times 10^{-6} \times 1.2 \times 10^{-8}) = 4.0 \times 10^{13}$ ions/cm.², while for the 100P, $N_p = 4.1 \times 10^{13}$ ions/cm.² For $\theta = .10$ the total concentration of adions plus adatoms is 3.7 or 3.8×10^{13} par-

ticles/cm.² From this we conclude that, for small concentrations of cesium on tungsten, nearly all the cesium is adsorbed as adions.

Most of the experiments for cesium on tungsten were performed with polycrystalline tungsten wires. In nearly all cases, the crystal size was small compared to the wire diameter. No attempt was made to separate the emission from various planes. Consequently, one can deduce only average values of φ_e versus θ . This will be indicated by a bar above the φ . A plot of $\bar{\varphi}_e$ versus θ is shown in FIGURE 4. From this plot, an average value of $\ell = 1.0 \times 10^{-8}$, and EQUATION 4 we can compute N_p versus θ . From the definition of θ it follows that $N_p + N_a = 3.8 \times 10^{14}\theta$. Hence N_a versus θ can be computed. The results are plotted in FIGURE 5. The vertical lines indicate our estimates of the uncertainties.

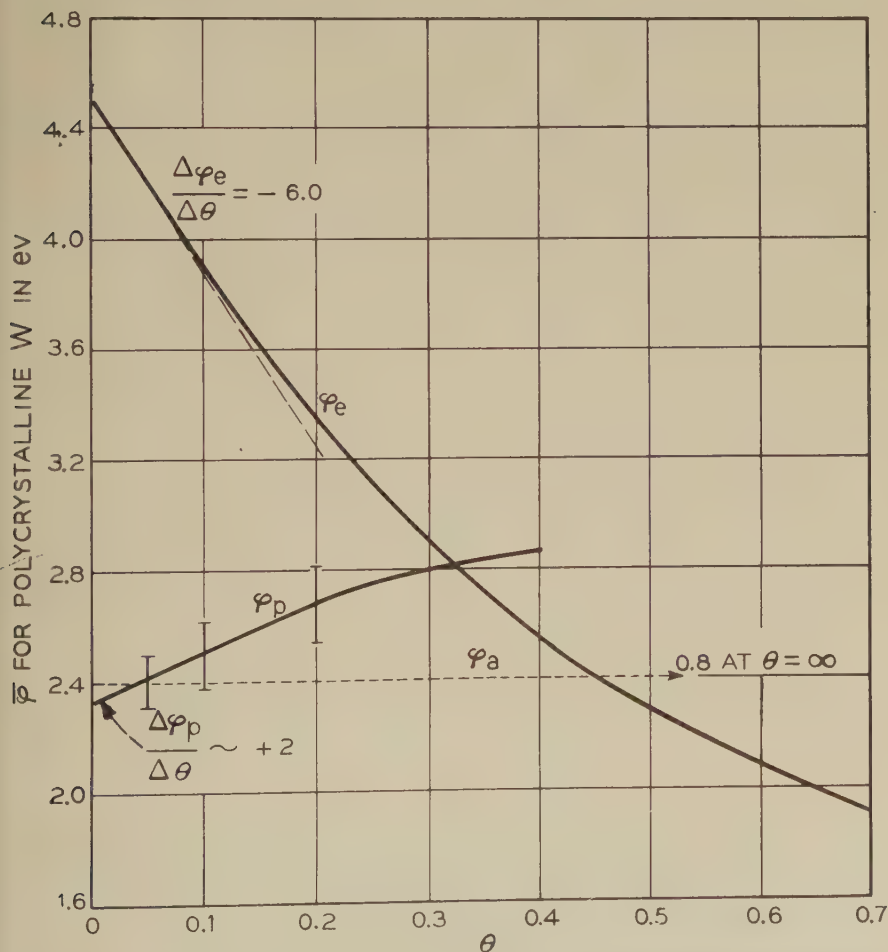


FIGURE 4. Average values of work function for electrons, atoms, and positive ions versus amount adsorbed: Cs on W.

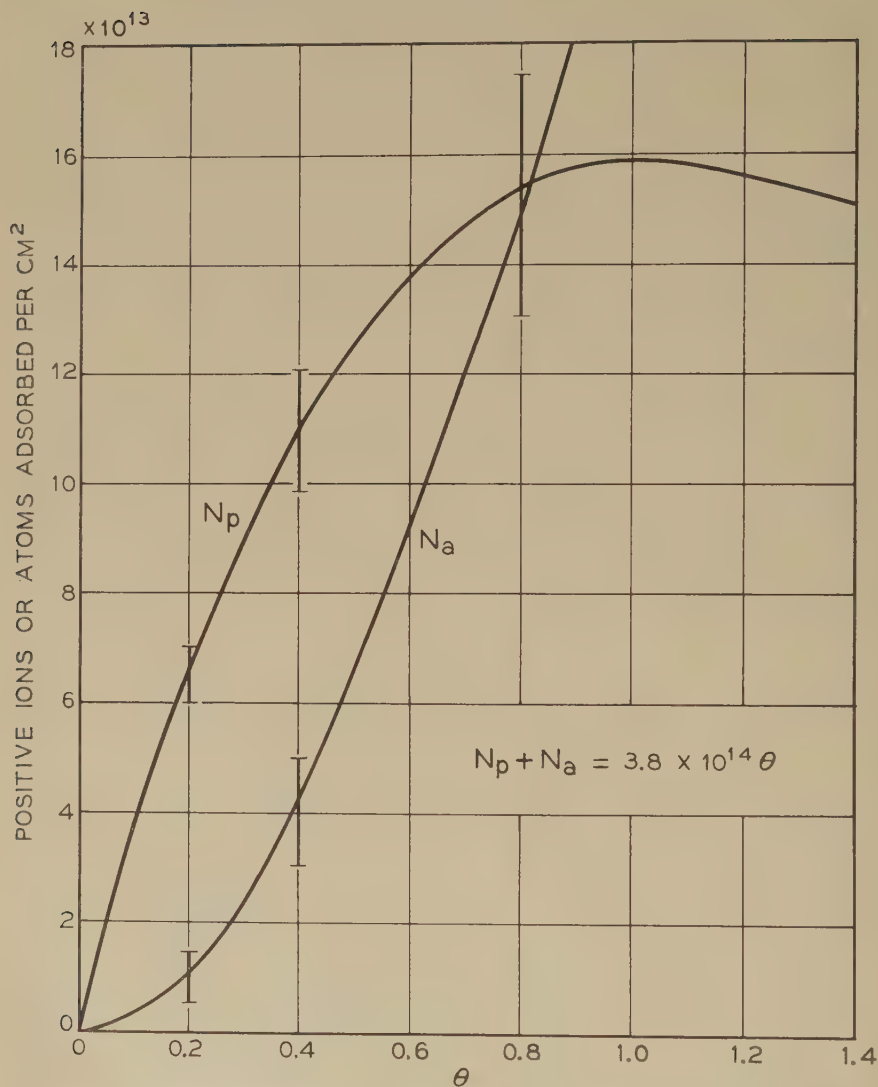


FIGURE 5. Concentration of adions (N_p) and adatoms (N_a) versus total amount of cesium on polycrystalline tungsten.

We are now ready to calculate average values of $\bar{\varphi}_p$ and $\bar{\varphi}_a$, the work necessary to evaporate adions and adatoms for various values of θ . These calculations are based on FIGURE 1, FIGURE 5, and EQUATIONS 2 and 3. The results are plotted in FIGURE 4. An estimate of the uncertainties is indicated by the vertical lines. At $\theta = 0$, $\bar{\varphi}_p$ and $\bar{\varphi}_a$ are nearly 2.4 volts while $\bar{\varphi}_e = 4.5$ volts. φ_a is practically independent of θ up to 0.7. However, for large θ , it must decrease since for bulk cesium $\varphi_a = 0.8$ volts. As θ increases near 0, $\bar{\varphi}_p$ in-

creases at the rate of 2 volts per layer, while $\bar{\varphi}_e$ decreases at the rate of 6 volts per layer. Owing to the uncertainties $\Delta\varphi_p/\Delta\theta$ might be 3 volts per layer, or half the rate at which φ_e decreases.

The adion theory not only explains why φ_e decreases with θ , it also explains why φ_p increases with θ and why $\Delta\varphi_p/\Delta\theta$ should only be a fraction of $\Delta\varphi_e/\Delta\theta$. To show this, it is necessary to consider an array of adions and their images, to choose a line normal to the surface, to compute the potential for a series of points on this line for all the dipoles, and to plot the potential versus distance from the surface. Zero distance is at the electronic image plane discussed above. FIGURE 6 shows an array of + signs such as might be found if adions were regularly spaced on the sites in a 110 plane. The distance between ions is such that θ is about 0.4. For ℓ we have used 1×10^{-8} cm. Curve C in this figure is an accurate plot of the potential versus distance above point C in the array. For a line above B, which is closer to an adion than is C, the potential rises more rapidly, goes above the potential for infinite distance, comes to a maximum, and then asymptotically approaches the same value as for curve C. For a line passing through an adion the potential rises to infinity at the distance ℓ , then decreases, and again asymptotically approaches the same large-distance value. In fact all lines have the same potential at large distances. This potential is the same as that which would be computed if the total charge were spread uniformly over a plane at distance ℓ above the electronic surface. In this case, the potential rises linearly from 0 to ℓ , attains its full value at distance ℓ , and has this same value for all distances larger than ℓ .

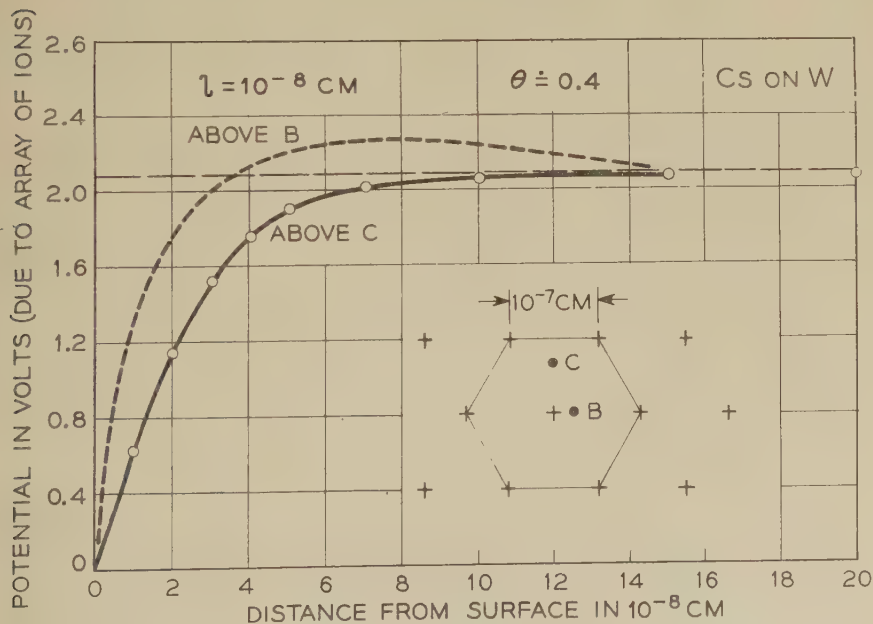


FIGURE 6. Potential due to a close-packed array of ions versus distance from the surface; $\ell = 10^{-8}$ cm.; $\theta = 0.4$.

If an electron escapes along a line above C, this potential field will help it escape. The work done on the electron in e.v. will be equal to the potential in volts shown in FIGURE 6. On the other hand, a positive ion escaping along the line C would be continuously pulled back toward the surface and hence the potential field makes it harder for the ion to escape. But the total work on the ion is only a fraction of the total work for the electron since the ion starts at distance ℓ while the electron starts at zero. For a line like B, the ratio of the ion work to the electron work is about 0.5, while for line C the ratio is 0.72. In these calculations, the location of the ions in the array is considered fixed. For an escaping electron, this is probably correct enough; but for an escaping ion this would leave a vacant ion position. At temperatures at which ions evaporate at appreciable rates, all adions are violently agitated and can move over the surface. Hence, as an adion is escaping from the surface, its previous neighbors move in to fill the vacancy. The effect of this movement is to decrease the ratio ion work to electron work. The calculated values of 0.7 to 0.5 might accordingly be reduced to 0.5 to 0.3, which are the experimental values.

We are now ready to consider an atom at a large distance from the nearly clean metal surface and plot a curve of the energy of the system as the atom approaches the surface along a line such as C in FIGURE 6. This type of curve is shown in FIGURE 7 for the curve marked Cs Atom. According to Lennard-Jones,¹⁰ such a curve approximately follows an inverse sixth power of the dis-

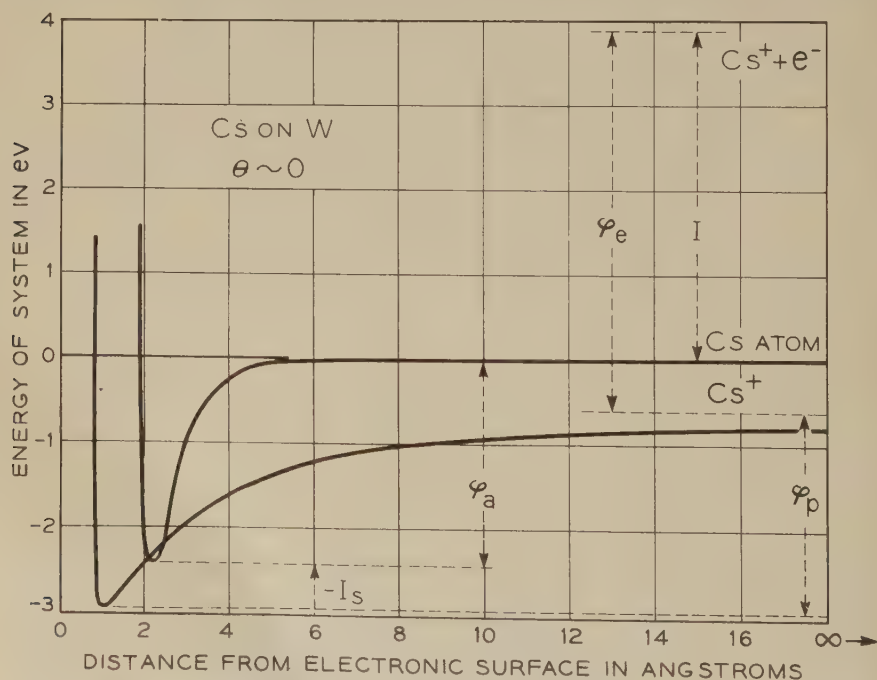


FIGURE 7. Potential energy of system versus distance from tungsten surface for an approaching atom or ion of cesium.

tance until it comes close to the equilibrium distance from the surface for an adatom, where very strong repulsive forces come into play. For still smaller distances, these repulsive forces are much larger than the attractive forces and the energy of the system rises approximately as the inverse twelfth power. The radius of a Cs atom is known to be about $2.7 \text{ \AA}$.¹³ Because of the topography of the tungsten atoms near the line C, the Cs atom can sink about $0.5 \text{ \AA}$. below the surface. The nucleus will accordingly be about $2.2 \text{ \AA}$. above the electronic surface. From what has been discussed above we also deduce that for this position, the energy of the system is 2.4 e.v. below zero. Hence, it is possible to plot the entire curve.

Suppose we again start with an atom at a large distance. Ionize the atom. This raises the energy by I or 3.9 e.v. Now let the electron drop into the Fermi level in the metal. This lowers the system energy by $\varphi_e \text{ e.v.}$ and leaves the Cs ion at a large distance. Now let this ion approach the metal surface. The energy decreases as the inverse first power in accordance with the image law. Near its equilibrium distance, strong repulsive forces again occur and the energy rises rapidly. The Cs ion radius is about $1.7 \text{ \AA}$. Because of surface topography the equilibrium ion distance is about $1.0 \text{ \AA}$.; and the system energy is $-(0.6 + 2.3)$ or -2.9 e.v. In this way, we obtain the Cs^+ curve in FIGURE 7. Other atom and ion curves could be drawn for lines other than those for line C.

It is now simple to derive an energy balance equation. By referring to FIGURE 7, the reader can verify that

$$\varphi_a = I_s + \varphi_p + \varphi_e - I \quad (5a)$$

or that

$$I_s = \varphi_a - \varphi_p - \varphi_e + I \quad (5b)$$

In this equation I_s is the surface ionization potential. It is defined as a positive quantity if work must be done to convert an adatom to an adion. It is negative if the energy of the system for an adion is lower than for an adatom. From EQUATION 5b and FIGURE 7 it follows that for Cs on nearly clean tungsten $I_s = -0.5 \text{ e.v.}$

In FIGURE 7, it is noteworthy that, at nearly all distances, the energy of the system is lower for the ion curve than for the atom curve. This means that a cesium particle held at almost any distance would be found in its ionic state more often than in its atomic state if we give the system a long enough time for electron transfers to occur. Only for distances near $2.2 \text{ \AA}$. are the energies nearly equal and the probability of finding the cesium as an adatom or as an adion are nearly equal. An adatom at this distance can readily give up the electron or electrons which it shares with tungsten atoms. Unless this electron is very quickly replaced, the ion will move closer to the surface. If this ion now receives one or two electrons from the metal, strong repulsive forces occur and the adatom moves farther from the surface. On the basis of this model, there should be a continual interchange between adatoms and adions. It is then natural to compute the ratio of adions N_p , to adatoms N_a , at any

instant. This is given by the Boltzmann equation

$$N_p/N_a = 10^{-I_s 5050/T} \quad (6)$$

In our case,

$$N_p/N_a = 10^{+2525/T} = 10^{3.06}$$

for $T = 820$. This means that there are about 1000 times as many adions as adatoms.

It is also interesting to compute the ratio of ions to atoms which evaporate. From EQUATIONS 2, 3, 6, and 5 it follows that

$$\frac{E_p}{E_a} = \frac{N_p}{N_a} 10^{-(\varphi_p - \varphi_a) 5050/T} = 10^{(\varphi_e - I) 5050/T} \quad (7)$$

It is noteworthy that E_p/E_a depends only on φ_e , I , and T and does not depend on N_p , N_a , φ_a and φ_p . This means that the ratio of ions to atoms which evaporate may be drastically different from the ratio of adions to adatoms. That E_p/E_a must depend only on $\varphi_e - I$, can be deduced from thermodynamical reasoning.

So far we have dwelt on the case where the surface is only sparsely covered with cesium *i.e.*, $\theta \doteq 0$. As θ increases, I stays constant, φ_e decreases rapidly, φ_p increases more slowly, and I_s decreases more slowly than $(\varphi_e - I)$. At some θ near 0.10, $(\varphi_e - I) = 0$ and $E_p = E_a$ even though most of the cesium exists as adions. At a still larger θ near 0.8, $N_p = N_a$. But by now $(\varphi_e - I)$ is so negative that practically all the cesium evaporates as atoms.

Many other properties of adions and adatoms might be discussed, such as surface mobility or the mean lifetime of an adsorbed cesium particle. For these properties the reader is referred to the earlier *Life History*. We shall now pass on to adsorbed molecules or admols as illustrated by the case of oxygen adsorbed on tungsten. The experimental procedures and more experimental details will be published elsewhere in the near future. We shall here review only the outstanding conclusions and concepts.

Oxygen Adsorbed on Tungsten

When oxygen or nitrogen is adsorbed on a clean tungsten surface, heat is evolved in an amount which corresponds to about 6 e.v. per molecule or about 140 k. cal. per gram molecule.¹ Since these amounts are as large or larger than the heats involved in many chemical reactions, it is generally agreed that the oxygen is chemisorbed. By this is meant that the O—O bond is broken and is replaced by two O—W bonds in which the oxygen atom shares electrons with 3, 4, or 5 tungsten atoms depending upon the plane on which it is adsorbed. As the amount of oxygen adsorbed increases, the heat of adsorption decreases and attains values near 0.1 e.v. The additional oxygen is then thought to be physically adsorbed and probably exists as admols. By this statement we mean that the O—O bond is much stronger than the bond between the atoms in the molecule and the substrate of chemisorbed oxygen; if any forces are exerted on one atom in the molecule, this atom pulls its neighbor with it and the molecule acts as a unit.

From another line of experiments it can be determined that only a fraction of the molecules that strike a surface become chemisorbed and stick to it.⁶ The remainder re-evaporate in a very short time. The fraction that sticks is called the sticking probability; it has also been called the condensation coefficient. For nitrogen on tungsten, it has been determined that the sticking probability at 300° K. is about 0.5 on some surfaces and may be as low as 0.1 on other surfaces of tungsten. This probably means that the sticking probability depends on the crystallographic plane. For any one surface, it maintains a constant value until θ reaches a value at which there is about one nitrogen atom for four tungsten atoms. We call this one layer and set $\theta = 1$. For $\theta > 1$ the sticking probability decreases rapidly until near $\theta = 2$ its value is only 10^{-4} . For oxygen on tungsten the behavior is similar, but this system has not been studied quantitatively.

In the field emission microscope, oxygen at pressures in the neighborhood of 10^{-7} mm. can be adsorbed on a clean tungsten filament. One small portion of this filament consists of a needle-like point. This point has a radius of curvature of about 2×10^{-5} cm. and consists of an approximately hemispherical surface of a single crystal.⁶ Consequently, all possible planes are exposed on this surface. When about 10,000 volts are applied between the filament and an anode, the electrical field at the point is about 50 million volts/cm. while elsewhere the field strength is much lower. Consequently, electrons are emitted or pulled out from the tungsten only at the point. The field current is much greater from those planes which have low work functions than from those planes which have high work functions. The electrons from the point travel in straight lines normal to the hemispherical surface and strike a fluorescent screen where they portray a highly magnified pattern of intensities which shows all the symmetries expected from a body-centered cubic crystal. From such a pattern of intensities, from the calculated field strength, and from the Fowler-Nordheim equation, one can compute that the work function of tungsten varies from about 4.2 to 5.5 volts. As oxygen is adsorbed on the point, the work function of some planes increases more rapidly than for others, and hence the emission decreases more rapidly on some planes than on others. If the sticking probability of oxygen on any or all planes were known as a function of θ , it would be possible to obtain a φ_e versus θ curve for any plane. Since the sticking probability is not known, a different procedure is followed.

The filament and point are flashed at about 2500° K. in a low residual pressure of oxygen to clean the surface. The filament is allowed to cool to room temperature. Oxygen is introduced until the pressure is 10^{-2} or 10^{-1} mm. After about 5 minutes, the oxygen is pumped out. By now, all metal and glass surfaces are heavily covered with adsorbed oxygen. The system is baked for one half hour at about 350° C., and allowed to cool. The pressure is then about 10^{-9} mm., which is so low that the residual gas does not interfere seriously with subsequent experiments. Voltage is now applied until the emission current is 1.0 microampere, and the pattern on the screen can be photographed in about one-fifth second. The voltage is removed and the point is heat treated at 400° K. for one minute. As soon as it has cooled to about 300° K., the voltage is increased until the field emission is 1.0 microamp. The pattern is

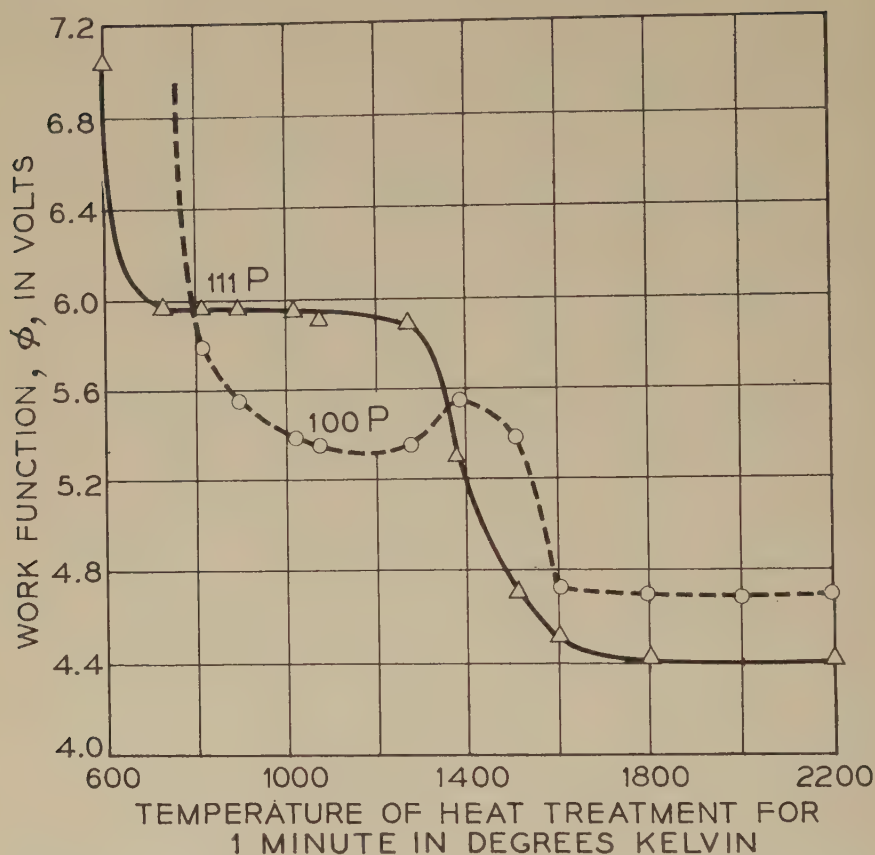


FIGURE 8. Variation of electron work function with temperature of heat treatment for the 111 and 100 planes of tungsten covered with oxygen.

photographed. This procedure is repeated for each 100° K. temperature interval up to 2200° K. From the density of the negatives and the observed voltages, it is possible to compute a work function θ_e for any plane after each heat treating temperature, T_h .

Until T_h is 600° K., the pattern consists of intensely bright units which bear little relation to the tungsten planes. They will be described and analyzed below. Above 600° K., the patterns show all the symmetries and plane locations shown by clean tungsten but the intensity distributions are different after each heat treatment. From these patterns, we have computed φ_e for the 111 plane as a function of T_h .

The results are plotted in FIGURE 8. Up to 600° K., φ_e is 7.0 volts or higher. After one minute at 700° K., φ_e has dropped to 6.0 volts. For heat treatments between 800 and 1300° K., φ_e is nearly constant at 5.95 volts. Then it drops rapidly to 5.2 volts for $T_h = 1400^\circ$ K. Above 1600° K., φ_e approaches a constant value of 4.4 volts, which is its value for the 111 plane for clean tungsten.

We interpret this curve as follows: Below 600° K., the 111 plane is covered with 2 layers or more of oxygen and the oxygen from the first 2 layers is not evaporated appreciably. At 700° K., the evaporation rate is quite rapid and most of the second layer is removed in one minute. For T_h between 800 and 1300° K. only one layer remains and its evaporation rate is so low that in one minute, θ changes inappreciably from 1.0. However, at 1400° K., the evaporation rate is quite rapid and θ decreases in one minute to perhaps 0.5. After one minute at 1800° K., the 111 plane is practically clean. From the temperatures at which φ_e drops rapidly, we deduce that the heat of adsorption for oxygen in the second layer is about 1.7 e.v., while in the first layer it is about 3.4 e.v.

The curve for the 100 plane is similar to that for the 111 plane but the temperatures at which φ_e drops rapidly are about 800° K. and 1600° K. From this circumstance, we deduce that on the 100 plane the heat of adsorption is about 2.0 volts for the second layer and 4.0 volts for the first layer.

By the first layer, we mean a concentration when every tungsten atom contacts one and only one oxygen atom. In the 111 plane, an oxygen atom contacts 3 and only 3 tungsten atoms; while in the 100 plane, it contacts 4 and only 4 tungsten atoms. We believe this to be the reason why the heat of adsorption on the 100 plane is about 20 per cent greater than for the 111 plane. In the first layer, we think of the oxygen making first valence bonds with the tungsten. By second layer, we mean that concentration when every tungsten atom contacts two and only two oxygen atoms; we think of the oxygen making second valence bonds with the tungsten.

Beyond two layers, there is little or no room left for an oxygen molecule to touch a tungsten atom, and by now the valence electrons of the tungsten atoms are nearly completely satisfied. There is accordingly little tendency for them to combine with any more oxygen, and there are no strong forces from the tungsten atoms to break the O—O bond. The valences of the adatoms of oxygen may not be completely satisfied and there may be some residual forces between these and molecules in the third layer. Accordingly, at temperatures near and slightly above 300° K. we should expect only a few oxygen molecules to remain adsorbed. These molecules should form protuberances on the surface at which the field strength should be unusually high. Since the field emission increases very rapidly with field strength, nearly all the emission should come from a relatively few highly localized molecules in the third layer. We might therefore expect the pattern to show little relation to the underlying planes.

This expectation is in fact realized in the patterns below 600° K. The emission comes from bright units. A unit may consist of a single spot, a doublet shaped like a barley corn, four bright spots at the corners of a square, a ring shape or more rarely, other shapes. These shapes are sketched in FIGURE 9. If one observes such a pattern at 300° K. for a matter of minutes, it is found that any one unit may *suddenly* change into another type or may disappear completely, or may drastically change its brightness. When one spot of a unit changes, all spots in that unit change. Sometimes a new unit will suddenly appear. These changes are indicated in the first two lines in FIGURE 9.

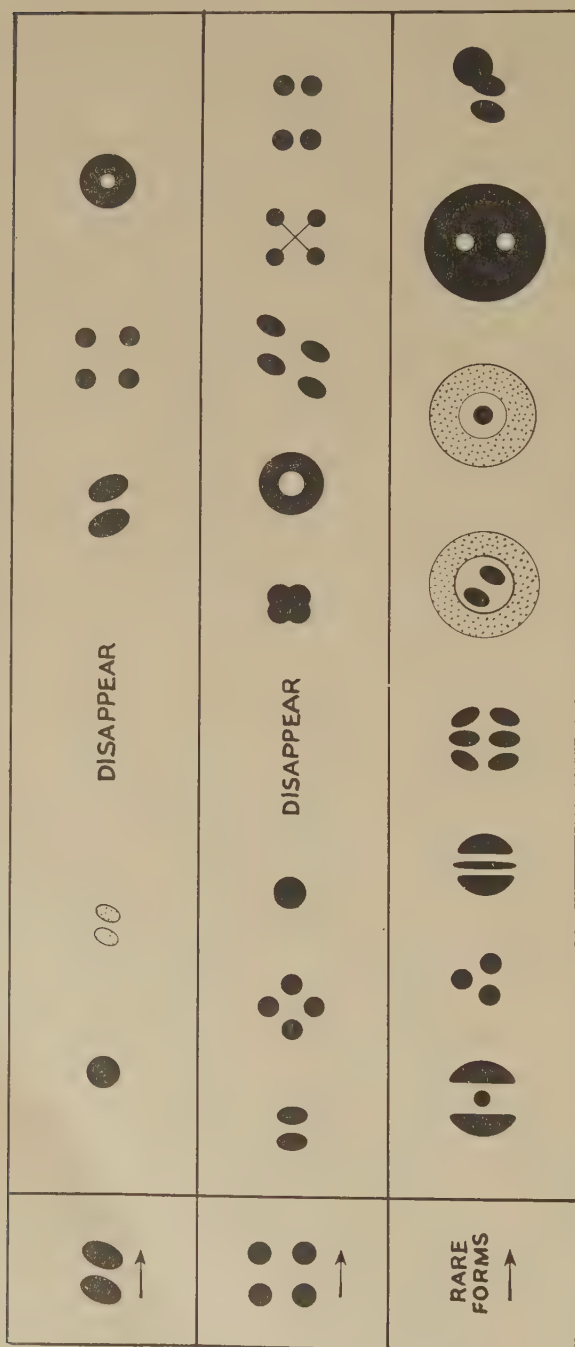
UNITS AND TRANSITIONS OBSERVED FOR O_2 ON O-W

FIGURE 9. Abrupt transformations observed for a doublet (top row); for a quadruplet (middle row); and rarer forms of units (bottom row). The last sketch in the bottom row shows the overlapping on the screen of a doublet and a singlet. The last sketch in the middle row shows an O_4 complex splitting into two O_2 complexes.

Heating the point slightly increases the rate at which changes occur. The rate also increases with current density and with pressure of the residual gas, suggesting that some of the changes are associated with ion bombardment. At 500° K., the rate of disappearance of units is greater than the rate of re-appearance and gradually the number of units decreases.

A theoretical study of the potential distribution of a protuberance consisting of dumbbell-shaped spheres indicates that the field will be greatest near the top and outer parts of the spheres. The path of an electron stream coming from these parts is such that greater magnification and greater resolving power result. All these facts, and others not mentioned here, lead us to believe that each bright spot in a unit is attributable to an individual atom in a molecule which rests on the chemisorbed layers. A single spot may be due to a molecule in which one atom dips down into the substrate while its partner protrudes above the substrate. A doublet may be due to a molecule which lies flat. A quadruplet may be due to an O_4 molecule. A ring shape may be due to a flat molecule in which the atoms rotate about a common axis. We are unable to think of any other mechanism which would account for the intensity of the spots, their shapes, the sudden changes in shape, the necessity for a substrate, and the fact that a group of spots act as a unit. If they are indeed attributable to individual, simple, small molecules, there is hope that molecular forces, adsorption, and catalytic phenomena will soon be better understood.

Summary

From the way in which the thermionic, photoelectric, or field emission of electrons from metals varies with the amount of a foreign species of atom or molecule adsorbed on the surface, it is possible to deduce whether the adsorbed particle exists as an atom, ion, or molecule. The system cesium on tungsten is especially revealing because with it one can measure the rate of emission of electrons, positive ions, and atoms as a function of the amount adsorbed. From these data, we conclude: (1) for a small amount of cesium, every cesium particle exists as an adion (adsorbed ion); (2) these adions produce strong fields which decrease the electron work function and increase the rate of electron emission; (3) these same fields increase the work function for positive ions so that the more cesium is adsorbed, the fewer ions are emitted; (4) as the amount of cesium increases, the ratio of adions/adatoms decreases until for a monolayer it is about 0.4; and (5) the electrostatic moment per ion depends considerably on the crystallographic plane; dense packed planes produce large moments while open structure planes produce small moments.

From the field emission data for oxygen on tungsten we conclude: (1) In the first "layer," oxygen atoms make first valence bonds with the tungsten and have a heat of adsorption of about 4 e.v.; (2) in the second "layer," oxygen atoms make second valence bonds with a heat of about 2 e.v.; (3) densely packed planes have lower heats than do planes with an open structure; and (4) in the third "layer," the surface forces are no longer strong enough to decompose the oxygen molecule. A few of these "admols" protrude above the first two layers; the

applied electric field is highly concentrated in the neighborhood of the atoms of each molecule. As a result, the field emission is concentrated in a relatively few intensely bright, highly localized groups of spots which act as units. We believe the units are single molecules and that the bright spots in a unit reveal the location of the atoms in the molecule.

References

1. ADAM, N. K. 1941. *The Physics and Chemistry of Surfaces*. Oxford Univ. Press. London, England.
2. BECKER, J. A. 1926. Thermionic and adsorption characteristics of cesium on tungsten and oxidized tungsten. *Phys. Rev.* **28**: 341-361.
3. BECKER, J. A. 1929. The life history of adsorbed atoms and ions. *Trans. Electrochem. Soc.* **55**: 153-175.
4. BECKER, J. A. 1932. The use of thermionics in the study of adsorption of vapours and gases. *Trans. Faraday Soc.* **28**: 148.
5. BECKER, J. A. 1951. The use of the field emission electron microscope in adsorption studies of W on W and Ba on W. *Bell System Tech. J.* **30**: 907.
6. BECKER, J. A. & C. D. HARTMAN. 1953. Field emission microscope and flash filament techniques for the study of structure and adsorption on metal surfaces. *J. Phys. Chem.* **57**: 153.
7. GWATHMEY, A. T. & A. F. BENTON. 1942. The reaction of gases on the surface of a single crystal of copper. I. Oxygen. *J. Phys. Chem.* **46**: 969-980.
8. IVES, H. E. 1924. Photoelectric properties of thin films of alkali metals. *Astrophys. J.* **60**: 209-230.
9. LANGMUIR, I. & K. H. KINGDON. 1924. Electron emission from adsorbed films on tungsten. *Phys. Rev.* **24**: 510.
10. LENNARD-JONES, J. E. 1932. Processes of adsorption and diffusion on solid surfaces. *Trans. Faraday Soc.* **28**: 333.
11. MARTIN, S. 1939. On the thermionic and adsorptive properties of the surfaces of a tungsten single crystal. *Phys. Rev.* **56**: 947.
12. MUELLER, E. W. 1953. Feldemission. *Ergeb. exakt. Naturw.* **27**: 290-360.
13. PAULING, L. 1945. *The Nature of the Chemical Bond*. Cornell Univ. Press. Ithaca, N. Y.
14. STRANSKI, I. N. & R. SUHRMANN. 1947. The electron emission of crystalline metal surfaces and its relation to the laws of crystal structure. *Ann. Physik.* **1**: 153; 169.

ADSORBENT-ADSORBATE INTERACTIONS AND SURFACE HETEROGENEITY IN PHYSICAL ADSORPTION

By J. M. Honig

Department of Chemistry, Purdue University, Lafayette, Ind.

Index

I. A. Scope of Review.....	742
B. Glossary of Symbols.....	743
II. Interactions Between Ideal Covalent Crystals and Nonpolar Gases.....	744
A. Forces between two isolated nonpolar atoms or molecules.....	744
B. Interactions between nonpolar gases and ideal covalent crystals; adsorption on uniform surfaces.....	746
(1) Calculation of dispersion energies by integration.....	747
(2) Calculations involving summations.....	750
C. The concept of surface heterogeneity in ideal crystals.....	753
D. Variation of potential minima along the surface; contour dia- grams and barriers restricting lateral mobility.....	753
III. Electrostatic Interactions Between (Polar) Gases and Perfect Ionic Crystals.....	755
A. The electric field and electric potential near the surface of an ionic lattice.....	755
B. Electrostatic interaction between ideal ionic lattices and gases..	758
(1) Gases characterized by the absence of permanent multipoles..	758
(2) Gases characterized by permanent multipoles.....	759
C. Polarization of the lattice by gases.....	760
D. Collaboration between forces; total adsorption energy.....	760
(1) General remarks.....	760
(2) Total adsorption energy; surface heterogeneity.....	761
E. Adsorption on metals.....	761
IV. Nonuniform Surface in Real Crystals.....	762
A. Effect of surface topography on adsorption energies.....	762
(1) Surface topography and dispersion potentials.....	762
(2) Surface topography and electrostatic potentials.....	763
(3) Impurities and surface heterogeneity.....	763
(4) Other effects.....	764
B. Experimental evidence relating to the existence of nonuniform surfaces.....	764
V. Distribution of Adsorption Energies Among Surface Sites.....	766
A. Mathematical formalism based on the Langmuir concept.....	766
(1) Determination of distribution functions from isotherm data; exact methods.....	769
(2) Determination of distribution functions from isotherm data; approximate methods.....	774
B. Critique of the inversion procedure.....	776

(1) Limitations attributable to assumptions implicit in the Langmuir theory.....	776
(2) Mathematical limitations of the method.....	780
VI. Gas Adsorption on Heterogeneous Surfaces in the Presence of Lateral Interactions.....	783
A. Fixed lateral interactions.....	783
(1) Adsorption on uniform surfaces.....	784
(2) Adsorption on heterogeneous surfaces.....	785
B. Variable lateral interactions.....	789
(1) Adsorption on uniform surfaces.....	789
(2) Adsorption on heterogeneous surfaces.....	790
C. Heterogeneity versus lateral interactions.....	791
VII. Surface Heterogeneity in Chemisorption.....	791
VIII. Bibliography.....	794

I A. Scope of review

This paper presents a review of current theories on the interaction between gases and solids, with special emphasis on the concept of surface heterogeneity. Owing to space limitations, the discussion will be confined principally to a consideration of processes in physical adsorption. Many related topics of great intrinsic interest, such as multilayer sorption or surface area determinations, are discussed only in so far as they contribute directly to the aim of the review.

The principal topics under discussion may be briefly summarized as follows: a description of van der Waals' interaction forces between a gas molecule and an isolated lattice atom is presented in Section II A, and is used as a basis for calculating the adsorption energies for gases on ideal surfaces in sections II B and II C. It is shown that the adsorption energy depends on the location of the adsorbate relative to the arrangement of the lattice atoms of the crystal. Methods for calculating contour diagrams that specify the variation of the adsorption energy over the surface are given in Section II D and may be used to compute the energy barriers impeding lateral mobility of the adsorbate.

Electrostatic interactions between adsorbate and adsorbent are taken up in Section III. Equations specifying the electric field and potential in the neighborhood of an ionic lattice are given (III A) and are used in later sections to obtain the electrostatic contributions to the adsorption energies.

Generalizing to the case of surfaces of real crystals (IV), it is shown that different topographical irregularities necessarily imply variations in the adsorption energies and hence surface heterogeneity (IV A). A review of experimental data confirming the presence of heterogeneity effects is presented in Section IV B.

Distribution functions characterizing the variation in adsorption energies over the surface are introduced, and mathematical procedures for their evaluation from isotherm data are discussed. The complexity of these theories depends on the number of factors affecting the adsorption process which one wishes to take into account. The mathematical formalism given in Section V A

can be used under the assumption that all sites associated with equal adsorption energy are covered in accord with the Langmuir theory. The very severe limitations of this formalistic method are discussed in great detail (V B), and it is emphasized that the procedure must be applied with great caution.

Section VI represents a review of theories of adsorption on nonuniform surfaces in which some of the restrictive assumptions of Section V are lifted. These theories are of much wider applicability but are mathematically rather untractable. Cases for which the general theory can be simplified are discussed (VI A and B).

The paper concludes with a brief discussion of the modifications which must be introduced when chemisorption phenomena are considered.

I B. Glossary of symbols

Whenever a given symbol is used in two or more senses the Roman numerals in brackets refer to the sections in which the symbol has different meanings.

a	Polarizability of atoms or molecules	k_1	Constants in EQUATION 6-23
a_2	Constant in the two dimensional van der Waals equation of state	k_2	
A	Surface area	K_i	Coefficient of variable parameters in EQUATION 2-8
b	The relation $\alpha(\epsilon, T)e^{-\epsilon/kT}$	$L(\eta)$	Function obtained from interpolation equation by replacing p in $\theta(p, T)$ with $\alpha_0/(e^{-\eta} - 1)$
b_2	Constant in the two-dimensional van der Waals equation of state	L (ν)	Fourier transform of L
c	Constant	m	Mass of particle
C_i	Constants in EQUATIONS 2-1 or 2-3 pertaining to the dispersion or repulsion energies between isolated atoms or molecules. Can be evaluated from physical properties of the interacting partners.	n	Index; number of sites of a given energy (VA)
d_0	Lattice parameter	p	Gas pressure
e	Base of natural logarithms; electronic charge	p_c	Pressure required for two dimensional condensation on a uniform surface
E_i	Adsorption energy on i th patch, inclusive of lateral interactions	p_i	Pressure p_c for the i th surface patch
E	Electrostatic Field	p_0	Saturation vapor pressure at the sample temperature
$f(\epsilon)$	Distribution density function for adsorption energies among surface sites, $f = dF(\epsilon)/d\epsilon$	q	Ratio r_0/z
$F(\epsilon)$	Fraction of surface sites with associated adsorption energies in the range 0 to ϵ	r	Distance between two molecules or atoms
$dF(\epsilon)$	Fraction of surface sites with an associated adsorption energy in the range ϵ to $\epsilon + d\epsilon$	r_0	Equilibrium value of r
$\left. \begin{matrix} g(b) \\ G(b) \\ dG(b) \end{matrix} \right\}$	Distribution functions defined analogous to $f(\epsilon)$, $F(\epsilon)$, and $dF(\epsilon)$, except that parameter involved is b rather than ϵ	s	Defined by EQUATIONS 6-8 or 6-11
h	Function defined in EQUATION 5-12	T	Temperature
H	Fourier transform of h (Section V A); heat of sorption (VII)	u	Quantity $\xi - \eta$
i	Running index; $\sqrt{-1}$ (V)	U_i	Interaction potentials between two isolated atoms or molecules attributable to the i th type of interaction force
$\left. \begin{matrix} j_a \\ j_L \\ j_s \end{matrix} \right\}$	Internal partition functions of atoms or molecules in the gas, liquid, or adsorbed phases, respectively	w	Lateral interaction potential between atoms or molecules adsorbed on adjacent sites on the surface
k	Boltzmann Constant	W_i	Interaction potential between adsorbent and adsorbate attributable to i th type of interaction force
		$\left. \begin{matrix} x \\ y \end{matrix} \right\}$	Coordinates in the plane of a smooth surface
		z	Coordinate perpendicular to plane of a surface
		z	Equilibrium value of z
		Z	Number of nearest neighbor sites

$\alpha(\epsilon, T)$	Collection of parameters defined in EQUATION 5-2	λ	Running index
α_0	$\alpha(0, T)$	μ	Running index (III A); permanent dipole moment (III B)
β	Defined in EQUATION 6-3	ν	Vibration frequency (II); integration variable for Fourier transforms (V A)
γ	Defined $2\pi i\nu$ (VA) or in EQUATION 6-18; angle (III B)	ξ	Defined by EQUATION 5-9
Δ	Change in quantity under consideration	π	3.14159 ...
ϵ	Adsorption energy	ρ	Number density of lattice atoms
ζ	Defined by e^u	σ	Variable of integration (II; see also FIGURE 2a); electron density in a volume element of the electron cloud (III B); defined in EQUATION 5-11; cross section of adsorbed molecule (VI B)
η	Defined by EQUATIONS, 3-8, 5-9, or 6-6	τ	Defined in EQUATION 3-2
θ	Fraction of the surface covered by adsorbate	ϕ	Electrostatic potential (III); spreading pressure (VI)
θ_i	Value of θ for the i th set of surface sites	φ	Defined in EQUATION 5-11
θ_e	Value of θ for all sites associated with an energy ϵ	Φ	Fourier transform of Φ
κ	Valence, multiplicity of electronic charge	χ	Relative pressure, p/p_0

II. Interactions between Ideal Covalent Crystals and Nonpolar Gases

This section deals with the interactions which are responsible for the physical adsorption of nonpolar gases on perfect covalent crystals. These interactions are characterized by the magnitude of the adsorption potential, W , which represents the energy required to remove an adsorbed molecule from the surface of a crystal to infinity. The adsorption potential W can be calculated (II B) once the interaction energy between two isolated atoms has been specified (II A), such interactions lead directly to the concept of surface heterogeneity for ideal, nonionic crystal lattices (II C).

II A. Forces between two isolated nonpolar atoms or molecules

The electron cloud configuration about the nuclei of atoms or molecules is associated with the presence of fluctuating dipoles or multipoles on these atoms or molecules. In the case of nonpolar gases these fluctuations are such that, on the average, the centers of positive and of negative charge coincide. The transitory multipoles will, however, induce similar asymmetries of charge on adjacent molecules. In addition, interactions occur with multipoles already present on neighboring partners. The assembly of multipoles just described, on the average, will remain in parallel orientation, since such a configuration leads to a maximum attractive interaction energy, statistically favored by the Boltzmann principle. Forces arising from this type of interaction are known as van der Waals dispersion forces.

This energy of attraction U between two isolated atoms or molecules was first evaluated by London,^{87, 88} who used the quantum mechanical perturbation theory. He obtained the approximate expression

$$(2-1) \quad U_d(r) = -\frac{C_1}{r^6} - \frac{C_2}{r^8} - \frac{C_3}{r^{10}}$$

where r is the distance between the two molecules and where the constants C_i depend on the physical properties of the interacting partners. The three terms

represent in order dipole-dipole, dipole-quadrupole, and quadrupole-quadrupole interactions.

The theoretical considerations leading to EQUATION 2-1 have been discussed in a review article by Margenau.⁹² Relations defining the constants C_i are also given in numerous other publications,^{19, 35, 92} making their further discussion unnecessary.

In numerical calculations the last two terms of EQUATION 2-1 are often neglected. Inspection of the C_i values tabulated for many common gases⁹² indicates, however, that neglect of the second term is justified only for $r \geq 8 \text{ \AA}$. This distance is considerably larger than the adsorbate-adsorbent distances normally prevailing in physical adsorption. Dipole-quadrupole interaction cannot be neglected in really accurate calculations of adsorption energies. The third term is generally quite small compared to the others, and *will* be neglected in subsequent calculations.

An equivalent derivation of an expression for U_d was obtained by Slater and Kirkwood,^{80, 121} using the variational principle.

Two molecules are also subject to a mutual force of repulsion which becomes prominent when appreciable interpenetration of their electron clouds occurs. Quantum mechanical calculations¹³ indicate that this repulsion potential is given by

$$(2-2) \quad U_r(r) = B e^{-r/r_0}$$

where B and r_0 are appropriate constants. Frequently an empirical expression of the form

$$(2-3) \quad U_r(r) = C_4/r^m$$

is used in place of (2-2). In this expression, the constant m is generally taken as 12 although values as low as 9 and 10 or occasionally as high as 20 or even 100 are needed for an adequate representation of repulsion energies.^{35, 84, 85, 141}

The total interaction potential is now given as a combination of EQUATION 2-1 with 2-2 or 2-3. For use in later discussions we shall adopt the relation

$$(2-4) \quad U(r) = -\frac{C_1}{r^6} - \frac{C_2}{r^8} + \frac{C_4}{r^{12}}$$

At a certain equilibrium distance, $r = r_0$, the forces of attraction and repulsion just balance; thus $(\partial U / \partial r)_{r=r_0} = 0$. Using EQUATION 2-4 in this condition one can write C_4 in terms of C_1 and C_2 as

$$(2-5) \quad C_4 = \frac{1}{2} C_1 r_0^6 + \frac{2}{3} C_2 r_0^4$$

Use of EQUATION 2-5 in 2-4 then yields

$$(2-6) \quad U(r) = -\frac{C_1}{r^6} - \frac{C_2}{r^8} + \frac{1}{2} \frac{C_1 r_0^6}{r^{12}} + \frac{2}{3} \frac{C_2 r_0^4}{r^{12}}$$

so that

$$(2-7) \quad U_0 \equiv U(r_0) = -\frac{1}{2} \frac{C_1}{r_0^6} - \frac{1}{3} \frac{C_2}{r_0^8} \quad (U_0 < 0)$$

If EQUATION 2-7 is now solved for $C_2 = C_2(C_1, U_0)$ and, if this value is used in EQUATION 2-6, one obtains finally

$$(2-8) \quad U(r) = -U_0[3(r_0/r)^8 - 2(r_0/r)^{12}] - \frac{1}{2}(C_1/r_0^6)[2(r_0/r)^6 - 3(r_0/r)^8 + (r_0/r)^{12}]$$

We note for future reference that if $C_2 = 0$, EQUATION 2-7 reduces to $U_0 = \frac{1}{2} \frac{C_1}{-r_0^6}$ and EQUATION 2-8 reads

$$(2-9) \quad U(r) = U_0[2(r_0/r)^6 - (r_0/r)^{12}]$$

A plot of U versus r , assuming (a) $C_2/C_1 = 3 \times 10^{-16}$ and (b) $C_2 = 0$ is presented in FIGURE 1. The depths of the minima in this diagram represent the magnitude of the attractive energy at the equilibrium distance r_0 .

II B. Interactions between nonpolar gases and ideal covalent crystals; adsorption on uniform surfaces

The dispersion and repulsion energies are known to be additive;^{88, 92} i.e., the total energy for a collection of molecules can be expressed as a sum of pairwise interactions over all pairs of the configuration. The energy of attraction holding an adsorbed molecule to the surface of a crystal can thus be calculated once the interactions $U_d(r)$ and $U_r(r)$ between the adsorbate and a single

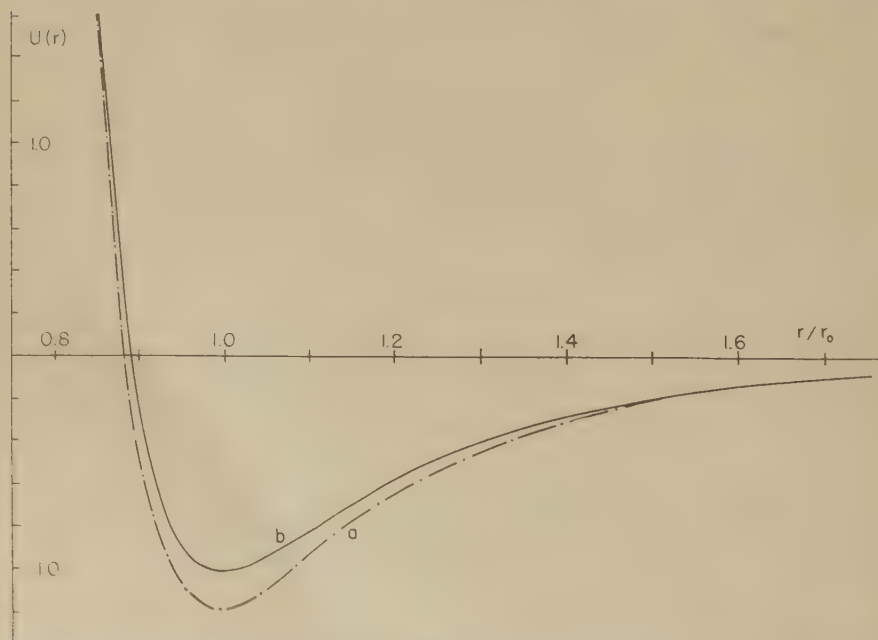


FIGURE 1. Van der Waals dispersion potentials, U , between a gas atom or molecule and an isolated lattice atom. Curve (a) $C_2/C_1 = 3 \times 10^{-16}$; (b) $C_2 = 0$; r is the distance of separation between interacting partners. U in k. cal. mole.

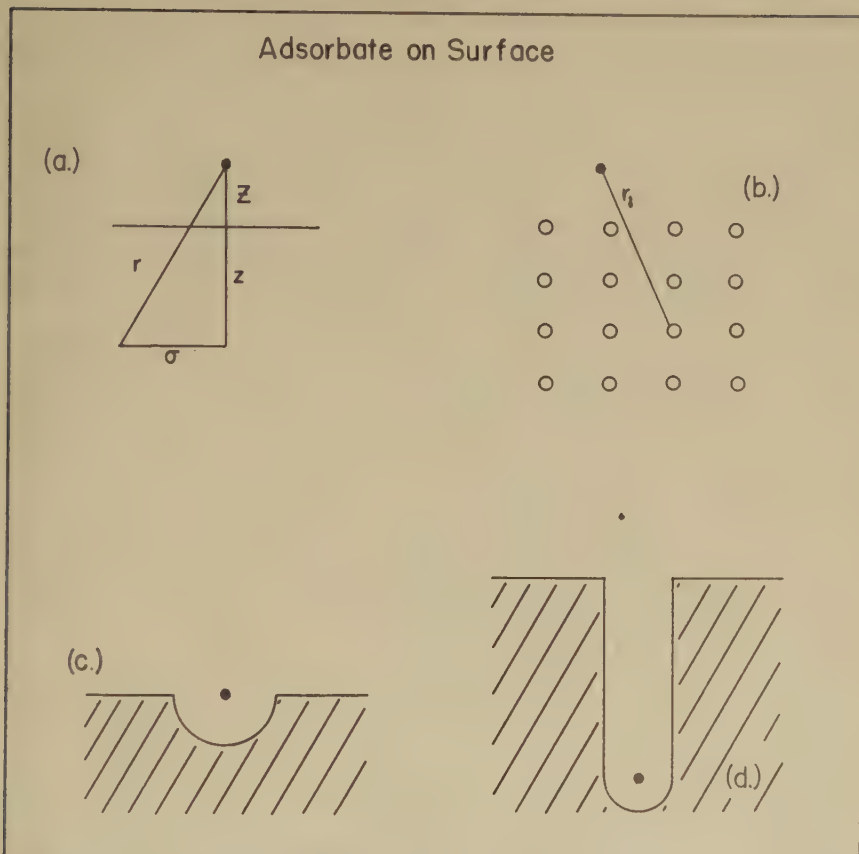


FIGURE 2. Adsorbate on surfaces of crystals. Adsorbate (a) on continuum bounded by plane surface, (b) on an ideal lattice bounded by smooth surface, (c) and (d) on surface of continuum characterized by various topographical features.

lattice atom are known. It is necessary to sum EQUATIONS 2-8 or 2-9 over all distances r_i between the location of the adsorbed particle and the lattice points of the perfect crystal (see FIGURE 2b; also ⁶⁹ for an illustration). A more detailed discussion of this process is deferred to section II B (2).

The summation procedure just mentioned can become quite tedious. Fortunately, it is possible to obtain qualitative information by use of an integration process described below.

II B (1). Calculation of dispersion energies by integration. In this method, the real lattice is replaced by a semi-infinite continuum with a mathematically plane bounding surface. The density of this equivalent continuum must be so adjusted as to match the mass of the k lattice atoms per unit volume of the crystal; then ρ represents the number of "equivalent lattice points" per unit volume of the continuum.

Let z be the shortest distance between the location of the gas atom and the

surface and let the particle be located at the origin of coordinates. The total interaction energy between the adsorbate and the semi-infinite solid is now

$$(2-10) \quad W(z) = \int_z^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty U(r) \rho \, dx \, dy \, dz$$

where x and y are orthogonal coordinates in the plane of the surface.

For simplicity, change to cylindrical coordinates, denoting by σ the radial distance from the z axis to a point in the interior of the continuum, so that $r^2 = z^2 + \sigma^2$ (see FIGURE 2a). According to EQUATION 2-8, $U(r)$ has the form $\sum_{j=1}^5 K_j r^{-n_j}$ where K_j is the coefficient of the j th term on the right-hand side of EQUATION 2-8 and where n_j assumes the values 6, 8, or 12. Thus

$$(2-11) \quad \begin{aligned} W(z) &= \sum_{j=1}^5 K_j \rho \int_z^\infty \int_0^\infty \frac{2\pi\sigma \, d\sigma \, dz}{[z^2 + \sigma^2]^{n_j/2}} \\ &= \sum_{j=1}^5 K_j \frac{\rho\pi}{n_j/2 - 1} \int_z^\infty \frac{dz}{z^{[2(n_j/2-1)]}} \\ &= \sum_{j=1}^5 K_j \left(\frac{\pi\rho}{(n_j - 3)(n_j/2 - 1)} \right) \left(\frac{1}{z^{[2(n_j/2-1)]-1}} \right) \end{aligned}$$

For $j = 3$, $n_j = 6$; for $j = 1, 4$, $n_j = 8$; and for $j = 2, 5$, $n_j = 12$. Substitution of these values in the last of the relations 2-11 leads to

$$(2-12) \quad \begin{aligned} W(z) &= \pi\rho r_0^3 \{ U_0(q^5/5 - 2q^9/45) \\ &\quad - (C_1/2r_0^6)(q^3/3 - q^5/5 + q^9/45) \} \end{aligned}$$

with $q = r_0/z$.

A plot of EQUATION 2-12 is given in FIGURE 3 for the hypothetical case $r_0 = 3.5 \text{ \AA}$. and $C_2/C_1 = 3 \times 10^{-16}$; thus $U_0 = -1.16(C_1/Zr_0^6)$. The curve passes through a minimum at $z = z_0 = 0.782r_0$; according to EQUATION 2-12 the depth of the minimum is

$$(2-13) \quad W_0 = W(z_0) = 0.458\pi\rho r_0^3 U_0 = -0.266\pi\rho C_1/r_0^3$$

To avoid confusion, it should be recalled that the quantities U_0 , r_0 , C_1 refer to the interaction between a gas molecule and an isolated lattice atom, while W , W_0 , z , and z_0 refer to the interaction between the gas molecule and the solid.

Nominally, the quantity $-W_0$ represents the energy required to remove the adsorbed atom back into the gas phase, *i.e.* the adsorption energy, while z_0 is the distance between the surface of the continuum and the adsorbate particle. However, several complicating factors must be considered: (a) The center of mass of the molecule vibrates in the potential well of FIGURE 3 along the z direction. Such vibrations are quantized and give rise to a set of discrete energy levels, which can be calculated as described by Hill^{67, 69} or by Orr.⁹⁶ The lowest vibration frequency was found by Hill to be of the order of magnitude of $\nu_z = 10^{12} \text{ sec}^{-1}$. This corresponds to a zero point vibrational level of approximately 100 cal./mole above the potential minimum W_0 . At 0° K .

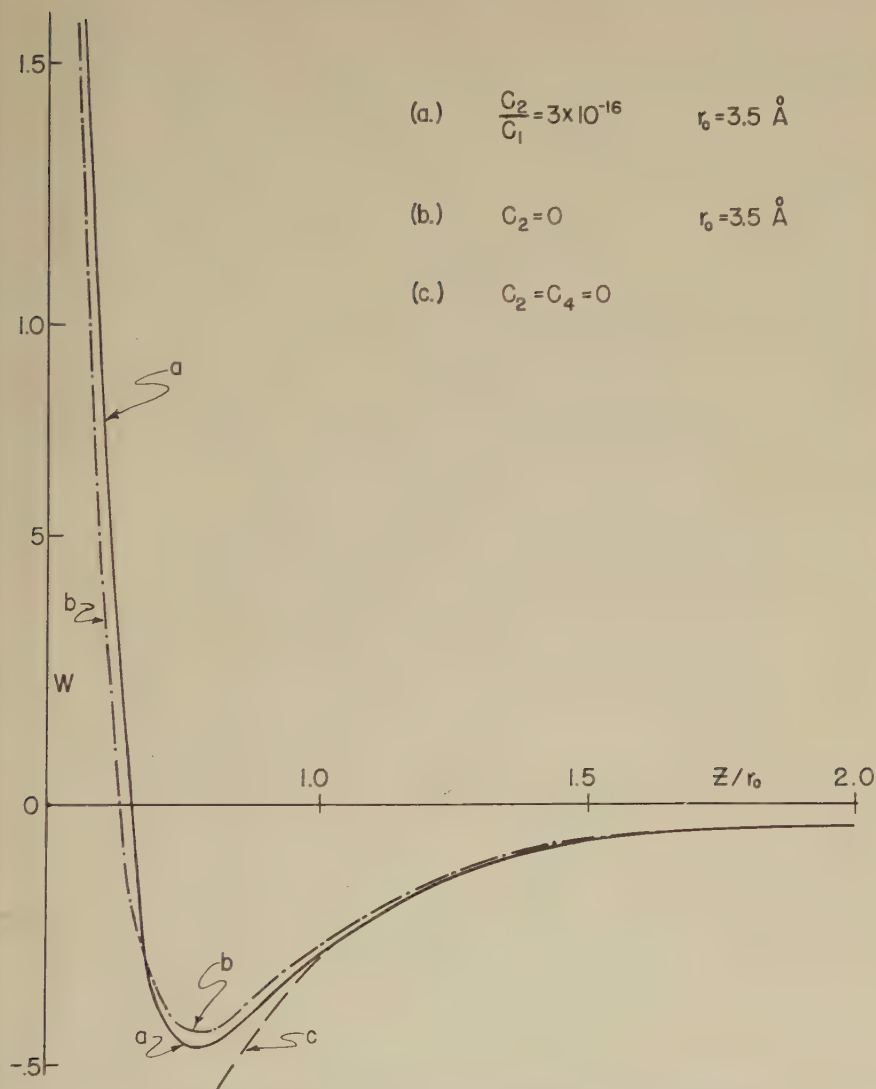


FIGURE 3. Van der Waals dispersion potentials for interaction between adsorbate and adsorbent continuum. Curves (a), (b), (c) obtained for various parameters; z is the perpendicular distance from the plane surface of the continuum. W in k. cal./mole.

the adsorption energy ϵ is thus given by $-(W_0 - hv_z/2)$. At higher temperatures, higher vibrational states may become populated, in which case ϵ is further diminished. (b) The adsorbed atoms interact with each other just as they do in the gas phase. This lateral interaction potential must be added algebraically to ϵ . At low surface coverage, and, in the absence of clustering, this correction is quite small. At intermediate coverage, the vibrational and

lateral potentials may partially compensate, in which case the quantity $-W_0$ is indicative of the adsorption energies to be expected from our crude model. (c) When a polyatomic molecule settles on the surface the adsorption energy depends also on its configuration with respect to the bounding plane of the continuum^{43, 67} (see below).

It is a consequence of the additivity law that the adsorption potential will be greatest if a maximum number of atoms in the molecule are in close contact with the surface. If other forces (Section III) constrain the molecule to be in different orientation the dispersion potential may be considerably diminished.

Various simplifications of EQUATION 2-12 are of interest. If dipole-quadrupole interactions are neglected by setting $C_2 = 0$, EQUATION 2-7 reads $U_{01} = -C_1/2r_{01}^6$ and EQUATION 2-12 simplifies to:

$$(2-14) \quad W_1(z) = \pi\rho_{01}^3 U_{01}(q^3/3 - q^9/45) \quad (U_{01} < 0)$$

identical with the equation derived by Hill.^{43, 67, 69} The subscripts 1 are used to differentiate the quantities r_0 , U_0 , and W used here from the somewhat different values found in EQUATION 2-12 where $C_2 \neq 0$.

A plot of EQUATION 2-14 is given in FIGURE 3b, assuming $r_0 = 3.5 \text{ \AA}$. The equilibrium distance z_{01} can be calculated by use of the condition $(\partial W_1/\partial z_1)_{z_1=z_{01}} = 0$, which leads to a value

$$(2-15) \quad z_{01} = (0.2)^{1/5} r_{01} = 0.765 r_{01}$$

The depth of the potential minimum, obtained by substitution of EQUATION 2-15 into 2-14 is

$$(2-16) \quad W_{01} \equiv W_1(z_{01}) = 0.497 r_{01}^3 \pi \rho U_{01} \quad (U_{01} < 0)$$

A final simplification can be effected by neglect of the repulsion potential (setting $C_4 = 0$). EQUATION 2-12 now reduces to

$$(2-17) \quad W_2(z) = -\pi\rho C_1/6z^3$$

a relation first derived by London⁸⁸ and discussed frequently in the literature.^{19, 32, 35, 99} A plot W_2 versus z is shown in FIGURE 2c. Under the assumed conditions, it reproduces the more complex equations quite well in the range $z \geq r_0$. Thus, at sufficiently large distances from the continuum, the attractive interaction potential 2-17 predominates. The latter falls off as z^{-3} ; by contrast, the interaction potential between isolated atoms diminishes as r^{-6} . The above calculations thus elucidate the nature of the long range forces operative between adsorbate and adsorbent; for further discussion of this point see Polanyi and London.⁹⁹

The use of the continuum model to represent a crystalline lattice is clearly quite unsatisfactory. This fact has been recognized in all papers on the subject; London⁸⁸ estimated that use of EQUATION 2-17 will yield sensible values only for $z \gg \rho^{-1/3}$, a condition hardly ever satisfied for gases adsorbed at the usual equilibrium distances of 3 to 5 $\text{\AA}$.

II B (2). Calculations involving summations. In order to arrive at a more satisfactory theory, it is necessary to take account of the crystal structure of

the solid and to compute the interaction between the adsorbate particle and the lattice atoms by direct summation. Such calculations have been carried out repeatedly^{4, 33, 83, 85, 86, 91, 141} under the assumption that the crystal lattice is perfect and that all the surface atoms lie in one plane. In many cases, a hybrid method is employed: The summation is carried out over all lattice points within a radial distance $r \gg \rho^{-1/3}$ of the location of the adsorbate particle and the contribution over the remainder of the crystal is then obtained by integration.

In all of these calculations, the dipole-quadrupole dispersion term C_2/r^8 has been neglected. As a result, the calculated interaction energies are somewhat low. However, it has been repeatedly pointed out that the repulsion energy roughly counterbalances the r^{-8} interaction potential.^{4, 18, 33, 35, 62, 83} If, accordingly, the term in r^{-8} is to be neglected in calculations, the repulsion term should be discarded likewise, as was done in a number of cases.

The adsorption energy, ϵ , obtained via W by use of the continuum model, is independent of the location of an atom on the surface. In the more refined calculations discussed below, ϵ is found to depend on the position of the adsorbate relative to the lattice structure of the solid. This is the case since the dispersion and repulsion potentials between two partners fall off so rapidly with distance. The magnitude of ϵ is thus controlled by the number of lattice atoms and their arrangement in the immediate vicinity of the adsorbate. For example, an adsorbed atom located on top of a lattice point (site d of FIGURE 4B) experiences a smaller interaction potential than one located above the center of a lattice cell (site a of FIGURE 4B) where it touches four surface atoms. Likewise, crystal faces characterized by different arrangements of surface atoms will be associated with different values of ϵ . Accordingly, ϵ or W are now functions of x , y , and z , but since their variation is periodic in x and y , it suffices to consider their variation over a unit cell of the crystal.

The above remarks can perhaps best be illustrated by use of a concrete example. For this purpose, we summarize the calculations of Orr⁹⁶ and Young¹⁵³ who calculated the adsorption potential of argon on the 100 and 111 faces of KCl. Using a fixed value of z , Orr computed the interaction energy for adsorption of A on sites (a) above the center of the lattice cell, (b) above the midpoint of the lattice edge, (c) above a K^+ ion, and (d) above a Cl^- ion (see FIGURE 4B). In these calculations, he summed the first term in EQUATION 2-9 over the 250 ions nearest to each of the four sites and obtained the remaining contribution by integration. The repulsive potential in 2-9 was summed over the nearest 14 to 16 ions. Contributions from the remainder of the crystal were found to be negligible. Since KCl is an ionic crystal, electrostatic interactions between adsorbent and adsorbate were also taken into consideration as discussed in Section III. However, these accounted for only 1 to 4 per cent of the total interaction energy. It may be assumed, therefore, that the effects described below are due solely to dispersion and repulsion forces. Orr repeated these calculations for several other selected values of z and fitted the results to an interpolation formula. In this manner, he constructed a plot of W versus z for each of the four locations of FIGURE 4B. These plots were similar to the curves depicted in FIGURE 3, but one important new factor emerged: the depth of the

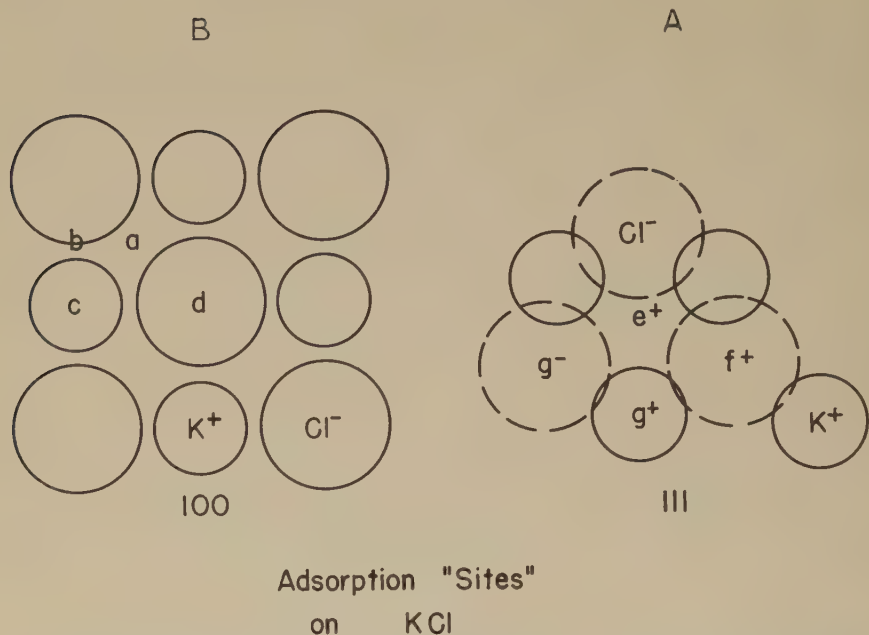


FIGURE 4. Adsorption "sites" for gas on 100 and 111 planes of KCl. Dotted circles on 111 plane diagram represent Cl^- ions located on next lower layer of the crystal.

potential minimum, W_0 , and the equilibrium distance, z_0 , were found to depend on the adsorption site considered. This effect is illustrated in TABLE 1.

In agreement with expectation, ϵ has its largest value above the center of the lattice cell. The smaller radius of the K^+ ion compared to the Cl^- ion accounts for the fact that ϵ is larger for site c than for site d.

The calculations by Young¹⁵³ follow those of Orr very closely. The 111 plane of KCl consists of triangular arrays of ions of the same sign (see FIGURE 4A). Young determined the interaction energy between argon and the lattice for the two types of triangular sites (e and f in FIGURE 4A) and for the location

TABLE 1
ADSORPTION ENERGIES AND EQUILIBRIUM DISTANCES FOR THE ADSORPTION OF A
ON KCl, AS A FUNCTION OF THE LOCATION OF THE ADSORBATE

100 Plane-Orr ⁹⁶				111 Plane-Young ¹⁵³			
Type of Site (See FIGURE 4A)	z_0 (Å)	$-W_0$	ϵ	Type of Site (See FIGURE 4B)	z_0 (Å)	$-W_0$	ϵ
		k. cal./mole				k. cal./mole	
a	3.22	1.62	1.59	e^+	2.80	2.10	2.06
b	3.44	1.34	1.31	e^-	2.61	1.47	1.45
c	3.48	1.45	1.42	f^+	2.78	2.24	2.21
d	3.74	1.27	1.23	f^-	2.67	1.36	1.34
				g^+	3.81	1.08	1.05
				g^-	3.77	0.70	0.68

g on top of an ion. For each type of site, one must distinguish between 111 planes made up of cations or anions.

Again the calculated curves of W' versus z resemble those of FIGURE 3, and, as in Orr's case, the position and depth of the potential minima varies with the type of site considered. The adsorption energies and equilibrium distances for the location of the potential minima are also entered in TABLE 1.

A study of this table shows that the adsorption energies depend (i) on the location of the adsorbent with respect to the lattice cell, and (ii) on the geometric arrangement of the surface atoms, *i.e.*, on the crystal face exposed to the gas phase.

The dependence of ϵ on these two factors has been verified elsewhere in the literature. Lenel⁸³ determined W'_0 for A adsorbed over the center of the lattice cell of the 100 and 110 planes of CsCl. Under the assumption that z_0 is given by the radius of A in the solid state and that the dipole-quadrupole dispersion energy cancels the repulsion energy, the calculated van der Waals dispersion energies on the 100 and 110 planes are 2.4 and 1.7 k. cal./mole respectively.

Crawford and Tompkins²⁵ estimated that the heat of adsorption of SO_2 on the 100 and 111 planes of BaF_2 differ by 1 k. cal./mole.

Barrer⁴ summed the dipole-dipole dispersion and the repulsion potentials for H_2 , A , and N_2 adsorbed at various locations on the basal surface of graphite. The potentials for adsorption over the center of each hexagon ring, above any apex, and above the bisector of an edge were found to be in the ratios 3.147:2.661:2.760.

A word of caution might be in order here: the correct choice of numerical values for the parameter z_0 is of critical importance in these calculations. Good apparent agreement between theory and experiment can often be achieved by a slight reshuffling of z_0 values. A critical appraisal of all z_0 values used in a calculation is necessary before it can be established whether agreement between theory and experiment is significant.³⁵

II C. The concept of surface heterogeneity in ideal crystals

The fluctuations of ϵ owing to the factors described in Section II B (2) can equally well be described as a distribution of ϵ values over the interface between a crystal and a gas. Whenever such a distribution or spread in ϵ values exists, such a surface is said to be heterogeneous or nonuniform. Again, the contributions of lateral interactions and of orientation effects are excluded from consideration.

II D. Variation of potential minima along the surface; contour diagrams and barriers restricting lateral mobility

The variations of W_0 with x and y can be most readily visualized by means of contour diagrams. These may be constructed by first preparing a series of plots of $W(z)$, as described in section II B (2), for a variety of locations (x_i , y_i) over the surface. From these plots, the depths of the potential minima can be read off as a function of x and y , and these values can be plotted on a contour diagram. Such diagrams were constructed for the adsorption of A on the 100 plane of KCl by Orr⁹⁶ (reproduced in FIGURE 5) and by Lennard-Jones.⁸⁵

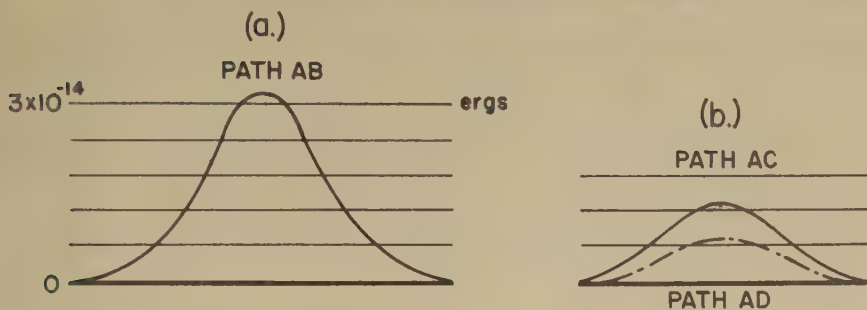


Contour Diagram for Dispersion and Repulsion Energies

FIGURE 5. Contour diagram for dispersion and repulsion energies of argon on 100 plane of KCl. Potential minima on "sites" A, B, C, D, etc. Potential maxima at E etc.

Inspection of FIGURE 5 reveals the presence of peaks above the Cl^- ions and of minima halfway between adjacent Cl^- ions, in agreement with the results of TABLE 1.

At sufficiently low temperature, an adsorbate particle will tend to be localized near the minima on the contour diagram (sites A, B, C, in FIGURE 5). Accordingly, it is permissible to identify the concept of "adsorption sites" with these minima in W_0 . This concept becomes elusive, however, for temperatures at which the particle acquires appreciable mobility; for, when the variations in the adsorption potential are small compared to kT , the adsorbate will move



Energy Barriers for Lateral Mobility

FIGURE 6. Potential barriers hindering translational motion of A on 100 plane of KCl along the paths AB, AC, AD of the preceding diagram.

without serious hindrance over the whole surface in an average field of force. The solid then *effectively* plays the same role as our continuum of Section II.B (1). Under these conditions the concept of an adsorption site loses its meaning.

Contour diagrams can be utilized to determine the energy barriers which restrict lateral mobility. It is obvious, for example, that a particle passing from A to B must overcome or tunnel through a barrier whose height is the difference $W_0(E) - W_0(A) = \Delta W_0$. A plot of the barrier height as a function of the path length is shown in FIGURE 6a. Barriers hindering lateral mobility for a variety of other paths are indicated in FIGURE 6b. It is seen that the heights range from 430 cal./mole to 86 cal./mole, compared to a value of $W_0(A) = 1630$ cal./mole.

In this particular case, the barriers restricting lateral mobility are very low along the paths of least resistance. At 43° K. the adsorbate possesses enough thermal energy, associated with degrees of freedom in the x and y directions, to overcome these barriers. Appreciable mobility will set in even at liquid hydrogen temperatures.

III. Electrostatic Interactions between (Polar) Gases and Perfect Ionic Crystals

If the adsorbate can be characterized by a permanent dipole or quadrupole moment, or if the adsorbent is made up of an ionic lattice, several new types of interactions between gas and solid are superposed on the dispersion forces discussed in Section II. These new forces will now be discussed.

III A. The electric field and electric potential near the surface of an ionic lattice

Fundamental to the later calculations is the determination of the electrostatic field $\mathbf{E}$ and of electrostatic potential ϕ at a given point outside an ionic lattice. Lennard-Jones and Dent⁸⁶ determined $\mathbf{E}$ and ϕ outside the 100 lattice plane of a simple cubic crystal by use of a method due to Madelung⁹⁰ and modified by Born.¹⁴ They arrived at an expression

$$(3-1) \quad \phi = \frac{2\kappa e}{d_0} \sum_{\substack{\lambda, \mu \\ \text{odd}}} \frac{(-1)^{(\lambda+\mu)/2}}{\sqrt{\lambda^2 + \mu^2}} \cos 2\pi \left(\frac{\lambda x}{2d_0} + \frac{\mu y}{2d_0} - \frac{\lambda + \mu}{4} \right) \frac{e^{-\pi z \sqrt{\lambda^2 + \mu^2}/d_0}}{1 - e^{-\pi \sqrt{\lambda^2 + \mu^2}}}$$

where κ is the charge on the lattice ion, $2 d_0$ the lattice parameter (edge of unit cell), and where the summation is to be carried out only for odd values of the indices λ and μ . The electric field normal to the 100 lattice plane can be obtained from the relation $\mathbf{E}_z = -\partial\phi/\partial z$. Considerable simplification is achieved by retaining only those terms for which $\lambda^2 + \mu^2 = 2$; the errors introduced by neglect of higher order terms are small.⁸⁶ One finds

$$(3-2) \quad \mathbf{E}_z = \frac{2\pi\kappa e}{d_0^2} \tau\left(\frac{x}{d_0}, \frac{y}{d_0}\right) e^{-\sqrt{2}\pi z/d_0} / (1 + e^{-\sqrt{2}\pi})$$

where

$$\tau\left(\frac{x}{d_0}, \frac{y}{d_0}\right) = \sum_{\substack{\lambda, \mu \\ (\lambda^2 + \mu^2 = 2)}} (-1)^{(\lambda + \mu)/2} \cos 2\pi\left(\frac{\lambda x}{2 d_0} + \frac{\mu y}{2 d_0} - \frac{\lambda + \mu}{4}\right)$$

The variation of $\mathbf{E}_z$ for various positions x and y over the lattice cell is specified by the function τ . It follows that $|\mathbf{E}_z|$ has a maximum at locations directly above any lattice ion and that $\mathbf{E}_z = 0$ on any plane normal to the surface which bisects a lattice edge. EQUATION 3-2 shows that $|\mathbf{E}_z|$ decreases exponentially with increasing distance from the surface. Numerical values of $|\mathbf{E}_z|e$ for various distances z/d_0 are shown in TABLE 2. Thus for NaCl ($d_0 = 2.82 \text{ \AA}$), $|\mathbf{E}_z|$ decreases by two orders of magnitude as z increases from 2.82 to 5.63 Å.

If the summation in (3-1) is restricted to the values $\lambda^2 + \mu^2 = 2$ it can be shown that

$$(3-3) \quad \phi = \frac{d_0}{\sqrt{2}\pi} \mathbf{E}_z$$

Values for $e\phi$ above a lattice point of NaCl are also given in TABLE 2.

As compared to the field associated with a point charge, $\kappa e/r^2$, the field due to crystal lattice is quite weak. This situation arises because every lattice ion is surrounded by nearest neighbors of opposite charge.

The variation of ϕ over the surface as specified by the function τ is again best visualized by a contour diagram such as shown in FIGURE 7. A positive charge would experience potential minima at positions above Cl^- ions. The potential barriers hindering lateral mobility along several paths are similar to those of FIGURE 6.

Several precautions must be observed in the application of the above equa-

TABLE 2
VARIATIONS OF $e|\mathbf{E}_z|$ AND OF $e\phi$ WITH z/d_0 ABOVE A LATTICE ION ON THE 100 PLANE OF NaCl.⁸⁷

z/d_0	0.500	0.584	0.600	0.667	0.700	0.750	0.800	0.883
$10^5 e \mathbf{E}_z $ (dynes)	8.44		3.46		1.42		0.548	
$-10^{13} e\phi$ (ergs)	5.36	2.45		1.23		0.64		0.28
z/d_0	0.900	0.916	1.000	1.5000				
$10^5 e \mathbf{E}_z $ (dynes)	0.240		0.0987	0.00186				
$-10^{13} e\phi$ (ergs)		0.14	0.06					

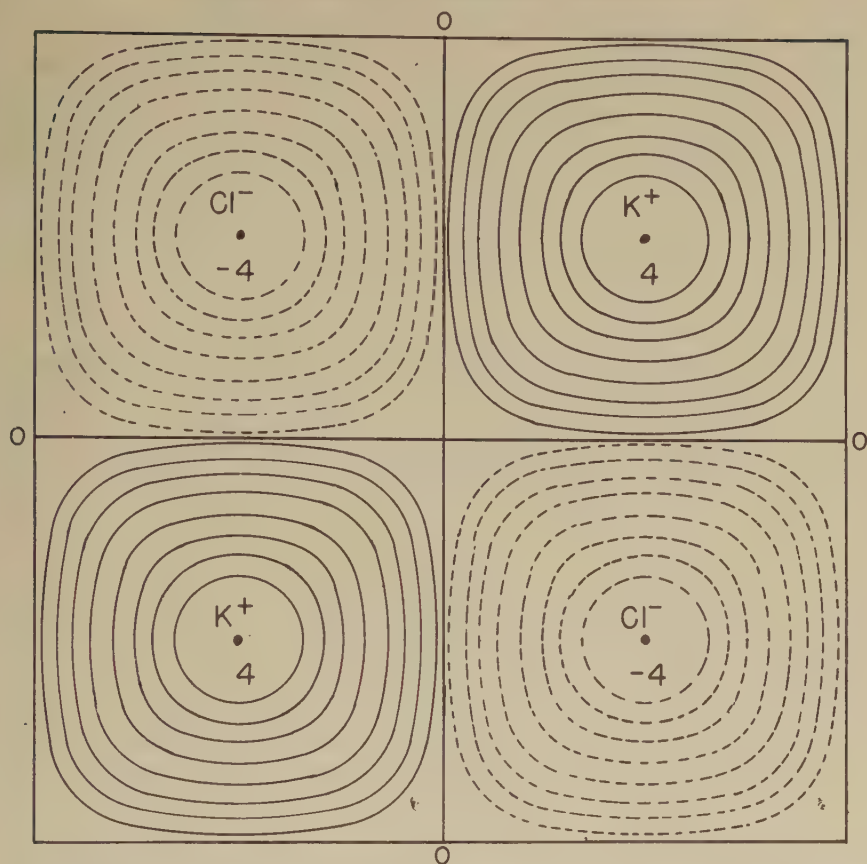


FIGURE 7. Contour diagram for electrostatic potential acting on a positive ion in the vicinity of the 100 plane of NaCl. Contour lines are in units of $0.4 (4\pi\epsilon/\sqrt{2}d_0) \{ \exp(-2\sqrt{2}\pi z/d_0)/[1 + \exp(-\sqrt{2}\pi)] \}$.

tions. It has been emphasized repeatedly in the literature^{106, 140} that the distance of the outermost layer of ions from the second layer is less than the distance between corresponding lattice planes in the interior of the crystal. Also, Verwey¹⁴² has concluded that in alkali halides the negative ions on the surface are displaced outward while the positive ions are displaced inward. This gives rise to a type of "staggering" of surface ions. Suitable refinements in the theory may have to be introduced to take these factors into account.

For positions above the lattice ions, EQUATION 3-2 simplifies considerably. Let the origin of coordinates coincide with the position of the center of mass of a positive lattice ion on the surface. Then in EQUATION 3-2, $x = y = 0$ and $\tau(0, 0) = 4$. Moreover, if we neglect $e^{-\sqrt{2}\pi}$ compared to unity, EQUATION 3-2 reduces to

$$E_z = \frac{8\pi e\kappa}{d_0^2} e^{-\sqrt{2}\pi z/d_0}$$

a relation cited by Hückel⁷² and by de Boer.³⁵ It may be shown that at these positions $\mathbf{E}_x = \mathbf{E}_y = 0$.

A number of other calculations similar to those of Lennard-Jones and Dent have been described in the literature. Lenel⁸³ has given formulae for the potential outside the 100 and 110 faces of body centered lattices such as CsCl. Orr⁹⁶ has computed the resultant electric field outside the 100 plane of a body-centered cubic lattice by summation over 88 to 96 ions nearest to the location of interest and by integration over the rest of the lattice. The field over the center of a lattice cell is in the opposite direction to that above the intersection of lattice planes. These calculations were extended by Young¹⁵³ to the 111 plane of simple cubic lattices. Here the field above the center of the triangular sites is in the opposite direction to that above the surface ion.

III B. Electrostatic interaction between ideal ionic lattices and gases

The electrostatic interaction between ideal ionic lattices and a gas atom or molecule will be discussed under two headings.

III B (1). Gases characterized by the absence of permanent multipoles. In this case one must take into account the polarization of the gas in the electric field of the lattice. In a uniform field of magnitude $\mathbf{E}$ the polarization energy is given by

$$(3-4) \quad W_p = -\frac{1}{2} a \mathbf{E}^2$$

where a is the polarizability of the molecule. This formula has been used in almost all calculations. As pointed out by Lenel,⁸³ however, the electric field outside ionic crystals is too inhomogeneous to permit the use of EQUATION 3-4. That portion of the adsorbate atoms which faces the surface will be much more polarized than portions further away from the surface. Accordingly, it is necessary to subdivide the total volume of the electronic cloud of the adsorbate into a number of elements dV and to determine the polarization effect in each element. In the limit of a fine subdivision, the polarization energy is:^{19, 35, 83}

$$(3-5) \quad W_p = - \int \frac{\phi_s - \phi_0}{I} \sigma dV$$

In this relation, ϕ and ϕ_0 are the electrostatic potentials in the volume element dV and at the center of the atom respectively, σ is the electron density in the volume element, and I is a "characteristic energy"—roughly the ionization potential of the atom. Detailed procedures for the calculations are given by Lenel.⁸³

Calculations cited by de Boer³⁵ give an indication as to the magnitude of these interaction energies. For the adsorption of argon on KCl, de Boer selected the values $a = 1.68 \times 10^{-24} \text{ cm}^3$ and $z_0 = 3.14 \text{ \AA}$. The electric field at this distance above a potassium ion is $\mathbf{E}_z = 1.45 \times 10^5 \text{ esu}$. Use of this value in EQUATION 3-4 leads to the value $W_p = -255 \text{ cal./mole}$. By contrast, use of EQUATION 3-5 yields $W_p = -450 \text{ cal./mole}$. Similarly, above the center of the lattice cell $\mathbf{E}_z = 0$; according to EQUATION 3-4, $W_p = 0$. Use of EQUATION 3-5 leads to a value of -370 cal./mole , however, because the polarization of those portions of the electron clouds not directly over the center of the cell is taken into account. The desirability of using EQUATION 3-5 in place of 3-4 for an accurate determination of polarization energies is evident from these examples. Fortunately, these

energies are usually small compared to the dispersion energies of Section II, so that use of EQUATION 3-4 in place of 3-5 will not lead to a serious error in the total adsorption energy.

III B (2). *Gases characterized by permanent multipoles.* In this case, other interaction energy terms arise which must be added to the dispersion and polarization potentials, $W_d + W_p$.

(a) If a molecule is characterized by a permanent dipole moment μ , the interaction energy with the field $\mathbf{E}$ is given by

$$(3-6) \quad W_m = -\mathbf{E} \cdot \mu$$

In the more accurate calculations, the inhomogeneity of the field over the distance of separation l of the dipole charges κe must be considered. If we assume that the dipole is aligned parallel to the z axis and that its center is located at z_0 we find

$$(3-7) \quad \begin{aligned} W_m &= -\kappa e \cdot \phi(z_0 - l/2) + \kappa e \cdot \phi(z_0 + l/2) \\ &= - \int_{z_0 + l/2}^{z_0 - l/2} \kappa e \mathbf{E}_z dz \end{aligned}$$

Numerical examples for this case are again supplied by de Boer³⁵ who considered the adsorption of OH dipoles on Cl^- ions at a KCl surface. Using 1.4 Å. as the radius of the OH group, $l = 1.0$ Å. as the distance of separation of O and H atoms, 1.8 Å. as the radius of the Cl^- ion, $\mu = 1.7D$, and $\mathbf{E}_z(z_0) = 2.15 \times 10^5$ esu, he obtained $W_m = -5.25$ k. cal./mole from EQUATION 3-6 and $W_m = -5.80$ k. cal./mole from EQUATION 3-7. These examples show that, for molecules or radicals with relatively large dipole moments, such as OH, the dipolar binding energy to surface ions may become larger than the contribution owing to the dispersion energy.

The calculations in connection with EQUATION 3-7 apply only to the case where the dipole axis is situated normal to the surface above a lattice ion. If at that location the axis were constrained to remain in parallel orientation to the surface, W_m would either vanish according to EQUATION 3-6 since $\mathbf{E}_x = \mathbf{E}_y = 0$, or remain small (EQUATION 3-7).

(b) Many molecules, such as CO_2 or N_2 do not possess dipole moments but can be characterized by quadrupole moments. Drain⁴² has shown that quadrupole interactions may contribute appreciably to the total adsorption energy. A complicating factor in such calculations is that the quadrupole moments for molecules are not known with certainty. A tentative tabulation of these moments is given in Drain's paper for a variety of simple molecules.

Molecules possessing a "linear" quadrupole moment Θ (see for example Brunauer,¹⁹ page 185) are subject to an interaction energy^{42, 72}

$$(3-8) \quad W_q = \frac{1}{2} \Theta \partial^2 \phi / \partial \eta^2 = \frac{1}{2} \Theta \partial \mathbf{E} / \partial \eta$$

in the external field of the ionic lattice. Here η represents a coordinate along the axis of symmetry of the quadrupole.

Drain⁴² computed the interaction energies W_q for positions a, b, c, d (FIGURE

4B) above the 100 plane of a cubic lattice. The quantity $\partial \mathbf{E} / \partial \eta$ was determined following the method of Lennard-Jones and Dent⁸⁶ and of Orr.⁹⁶ For N_2 on KCl (with z_0 given in TABLE 1) the maximum $-W_q$ values corresponding to positions a, b, c, d were 320, 870, and 1150 and 400 cal./mole respectively. These maximum values are possible only for proper orientation of the N—N internuclear axis with respect to the lattice cell. For example, at the center of the lattice cell, the axis should be parallel to the surface diagonal connecting K^+ ions, while above the midpoint of the lattice edge, the axis should be at a 45° angle with respect to the surface and so oriented that its projection coincides with the lattice edge. Above K^+ ions the N_2 molecule should be aligned normal to the surface. As is seen from this work, the quadrupole interaction energies are not necessarily negligible.

In earlier calculations Lenel⁸³ estimated the charge distribution of the CO_2 quadrupole from experimental heats of adsorption on KCl. The contribution of the dispersion forces was subtracted, Θ was estimated and the electrostatic interaction with other ionic adsorbents was computed. Iliin^{73, 74, 19} derived an expression for W_q based on the assumption that the lattice field $\mathbf{E}_s$ falls off as $z^{-\xi}$, rather than exponentially, where ξ is an appropriate constant. This theory also utilizes macroscopic properties of the adsorbate, and for these reasons is not nearly as satisfactory as that of Drain.

III C. Polarization of the lattice by gases

In addition to the above interactions, the polarization of the lattice by gases with permanent multipoles ought to be taken into consideration. Generally, this process results in a small contribution which is often neglected. For a discussion of the difficulties attendant to exact calculations of this type see Lennard-Jones and Dent.⁸⁶

III D. Collaboration between forces; total adsorption energy

III D (1). General remarks. As is seen from the preceding sections, various forces may act jointly in holding adsorbate molecules to surfaces of ionic crystals. This collaboration has been studied for many cases by de Boer³⁵ to whom the reader is referred for further details. A few pertinent conclusions are briefly summarized below (see also de Boer and Custers³²):

(1) Dispersion forces tend to aggregate gas molecules in such locations and positions [compare Section II B(2)] as to provide for contact with the maximum number of surface atoms or ions. By contrast, electrostatic forces tend to localize the adsorbate particles at positions directly above lattice ions, and at orientations favorable for maximizing the electrostatic energies. Such orientations usually run counter to those required for maximizing the dispersion potential [Compare Sections II B(1) and III B(2)].

(2) Generally, in the adsorption of nonpolar gases on ionic surfaces, the dispersion potential outweighs the polarization potential.

(3) Molecules associated with very strong dipole or quadrupole moments may experience electrostatic interaction potentials as large as, or greater than, dispersion potentials. In that case, the location and orientation of adsorbate

molecules may be intermediate to those prevailing, if either type of force alone predominated.

(4) Complex molecules with peripheral dipoles may be oriented in special positions. For example, in the adsorption of phenol, the benzene ring will lie flat on the ionic surface, with the hydroxyl group oriented toward a negative surface ion.

III D (2). Total adsorption energy; surface heterogeneity. The total adsorption for gases on ionic crystals can be determined by the method described in Section II B (2). The only difference in procedure is that the contributions arising from the electrostatic interactions must be included in the resultant plot of $W(z)$ versus z . The calculation of vibrational energy levels and of adsorption energies is carried out as indicated by Orr.⁹⁶ Due account must be taken of the orientation of the adsorbate with respect to the surface.

Once the potential energy diagrams have been determined for a variety of positions, contour diagrams may be constructed similar to those described in Section II D. Adsorption "sites" on this diagram can then be located, and energy barriers hindering lateral mobility may be computed.

Again, bounding surfaces made up from different lattice planes will in general be characterized by different potential minima. On a given lattice plane, different locations are also associated with different potential minima. These differences in minima describe a set of heterogeneous surfaces.

III E. Adsorption on metals

We conclude with a brief summary of the adsorption of gases by metals. This subject has been reviewed extensively;^{19, 37} the reader is referred to these papers for further details.

Gases are thought to be held to metals by "image forces"; that is, any asymmetry of the charge distribution in the gas, whether temporary or permanent, will produce a corresponding image in the metal and thus give rise to attractive interaction forces.

The interaction between metals and gases with fluctuating charge asymmetry has been treated by Lennard-Jones,⁸⁵ Bardeen,³ and by Margenau and Pollard.⁹³ These authors, using various stages of refinement, found interaction potentials that vary as z^{-3} and that are independent of the surface location x and y . Prosen, Sachs, and Teller¹⁰² and Prosen and Sachs¹⁰³ pointed out that the above theories are valid only for large z . These authors derived expressions in which the dispersion potentials vary as z^{-1} for a free electron gas in the metal and as z^{-2} for a degenerate Fermi electron gas. Again, their expressions do not involve the structure of the underlying surface.

When polar gases are used as adsorbates the interaction with image forces can be treated by classical methods. The problem was studied by Lorenz and Landé⁸⁹ and by Magnus⁹¹ for gases with permanent dipoles. The interaction energy is given by

$$(3-9) \quad W_i = -\frac{\mu^2}{16z_0^3} (1 + \cos^2 \gamma)$$

where μ is the dipole moment and γ is the angle between the dipole axis and the normal to the surface. This contribution must be added to the dispersion energies described in the last paragraph. Values obtained from EQUATION 3-9 are quite low. Thus, for an OH group with $\mu = 1.7D$, $z_0 = 2.7 \text{ \AA.}$, and $\gamma = 0$ one finds $W_i = -265 \text{ cal./mole.}$ This value is considerably smaller than that obtained for the interaction of a dipole in the field of an ionic lattice [Section III B (2)].

Jacquet⁷⁵ introduced a refinement by taking account of the polarization of the adsorbate dipole by its own image in the metal. He deduced the relation

$$(3-10) \quad W_{ip} = -\frac{\mu^2}{8z^3} \left[\frac{2}{1 - \frac{a}{4z^3}} \cos^2 \gamma + \frac{1}{1 - \frac{a}{8z^3}} \sin^2 \gamma \right]$$

where a is the polarizability of the adsorbate. This equation leads to values that are 30 to 40 per cent larger than those of EQUATION 3-9. The interaction between molecules possessing permanent quadrupoles and metals was also studied by Jacquet.

IV. Nonuniform Surfaces in Real Crystals

So far, the energetics of adsorption on ideal crystals has been discussed. We consider next the modifications which must be introduced to describe the adsorption of gases on real crystals. Three complicating factors must be taken into consideration:

(a) Real crystals are finite in extent, so that a certain percentage of constituent atoms are located on edges or corners of the lattice network. Effects attributable to the presence of such atoms will be designated as "edge effects." These effects become of increasing importance with the diminishing size of the adsorbent particles.

(b) The individual crystallites are riddled with all types of lattice imperfections, fissures, cracks, elevations, and lattice dislocations. In short, the regular arrangement of atoms or ions in ideal crystal lattices is disturbed.

(c) Impurities, consisting of foreign atoms or of ions with abnormal valencies may be present, and/or vacancies resulting in lattice defect structures may be encountered. The nature of the lattice imperfections has been described in many papers.^{21, 50, 81, 98, 106, 115, 144}

IV A. Effect of surface topography on adsorption energies. All of the above lattice imperfections affect the adsorbent-adsorbate interaction energy. An adequate description of these effects cannot be given; first, because every type of lattice imperfection affects the interaction potential differently; second, the dispersion and electrostatic interactions are quite often altered in opposite directions by the introduction of a given type of lattice disturbance; third, the total interaction depends on the topographic distribution of the atoms associated with a given type of lattice imperfection or edge effect.

In view of the complications it is necessary to restrict oneself to semiquantitative arguments designed to show that any unevenness in the surface of a solid necessarily implies the existence of surface heterogeneity as defined in Section II C.

IV A (1). Surface topography and dispersion potentials. The relation be-

tween surface topography and dispersion potentials was studied by de Boer and Custers.³² These authors calculated W_d for a molecule adsorbed at the center of a variety of lattice imperfections. In their calculations, they used the integration procedure discussed in Section II B (1). For example, a molecule adsorbed at the center of a hemispherical cavity of radius z_0 (see FIGURE 2c) is subject to a dispersion potential four times as great as for a molecule located on a smooth surface. Inside a long, narrow capillary, terminated by a hemisphere at whose center the molecule is located, the adsorption potential is roughly 6.7 that of the plane surface (FIGURE 2d). The largest potential is obtained for a molecule adsorbed in a spherical cavity of radius z_0 . This value is 8 times as large as that for a molecule at a distance z_0 from a plane surface. By contrast, molecules located at the top of very long and narrow peaks projecting out from the surface, experience potentials that approach the dispersion energies between the adsorbate and an isolated lattice atom. An overall spread of one order of magnitude in ϵ can thus be accounted for on the basis of the surface topography.

Using the summation method, de Boer and Custers repeated their calculations. They obtained, at least for the case of a simple cubic lattice, similar increases in dispersion potentials for the types of lattice imperfections considered above. This result is quite in accord with physical intuition. Inside a topographic depression an adsorbate particle is in close contact with a larger number of lattice atoms than on a smooth surface. Conversely, in elevated portions, the number of close-by lattice atoms is smaller. Since the dispersion potential depends primarily on the arrangement and the number of nearby lattice atoms [see Section II B (2)], any unevenness in the surface implies the existence of heterogeneity effects.

IV A (2). Surface topography and electrostatic potentials. The electrostatic potential outside an ionic lattice is also affected by lattice imperfections. Using EQUATION 3-3, de Boer found the following values of ϕ in units of e^2/z_0 : (a) on an ideal surface $\phi = -0.0662$; (b) at the edge of a cubic crystal $\phi = -0.0903$; (c) at a corner $\phi = -0.2470$; (d) adjacent to a step deformity $\phi = -0.8737$; (e) over a lattice ion in a rhombic dodecahedron face of an alkali halide $\phi = -0.2082$; and (f) in the limit of an isolated surface ion $\phi = -1.000$.

Stranski^{19, 123} has carried out a similar calculation without giving mathematical details. His numerical values differ from those cited above. Nevertheless, he notes variations of ϕ in the range from $-0.132 e^2/z_0$ to $-0.807 e^2/z_0$, depending on the location of the adsorbate with respect to lattice deformities or edges.

The physical basis of these results is again quite readily visualized. On a uniform surface, the regular alternation between positive and negative ions results in a partial cancellation of the field. The more nearly isolated a given lattice ion is from others (as on the peak of an elevation) the less is this cancellation effect in evidence.

The importance of edge or topographical effects in altering the electrostatic potentials has been reviewed repeatedly in the literature.^{19, 34, 51, 100, 140}

IV A (3). Impurities and surface heterogeneity. The effect of impurities differ with different cases. If the impurity is nonionic and located interstitially

it will contribute an independent term of the type (2-4) to the total interaction energy. If it replaces a lattice atom, the constants C_i in EQUATION 2-4 will differ with each impurity. The overall effect will also depend on such factors as the distribution and the location of impurity centers on the surface.

If the impurity is ionic and well separated from others on an essentially covalent lattice then the impurity represents a strong center of attraction for gas molecules with permanent multipoles. The same is true if, in an ionic lattice, the impurity consists of an ion of abnormal charge. Finally, the presence of a neutral particle in an ionic lattice will lead to a distortion of the potential in the neighborhood. All of these effects give rise to variations in ϵ over the surface and thus contribute to the nonuniformity in the adsorption energy.

IV A (4). Other effects. The heterogeneity of a surface may be altered by the presence of many crystallites which expose different crystallographic planes to the gas phase.

Again, heterogeneity effects may be owing to the presence of lattice defects. For example, in many metallic oxides, oxygen atoms are missing from surface sites, thus exposing metal atoms in the next lower layer of the crystal (see also Beech¹²).

The discussion of Sections IV A (1), II B (2), and IV A (2) shows that, when dispersion forces predominate, the adsorbate is bound most tightly in depressions, pits, fissures, cracks, or over centers of lattice cells. On the other hand, if electrostatic interactions predominate, the adsorbate tends to be located on elevations or above lattice points. Similarly, lateral interactions between two adsorbed molecules are always attractive when dispersion forces are dominant. Repulsive lateral interactions may occur when atoms with permanent multipoles are adsorbed side by side so as to form an electric double layer on the surface. The "antagonistic" tendencies of the two types of adsorption forces has been emphasized repeatedly in the literature.^{5, 32, 35, 36} It follows that the degree and type of nonuniformity of a given surface will appear to change with the type of gas used as the adsorbate.

IV B. Experimental evidence relating to the existence of nonuniform surfaces. It is difficult to cite clear-cut experimental evidence which proves the existence of energetically heterogeneous surfaces. However, much of the experimental work in the last two decades indicates that surface heterogeneity influences physical adsorption processes.

Use of the Mueller field emission microscope^{2, 7, 21} has provided one tool for such studies. In a review article by Ashworth,² the author presents a number of photographs showing preferential adsorption of gases by particular crystal faces of metallic probes. As the temperature is raised, the adsorbate is seen to evaporate from some of the crystal faces much more readily than from others. Unfortunately, at the temperatures of operation, the gas was probably held to the surface as a chemisorbed layer, and any deductions in terms of physical adsorption phenomena are clearly tenuous. Nevertheless, the fact that the adsorbate evaporates consecutively from different crystal faces on raising the temperature and that it settles back on the surface in re-

verse order on lowering of the temperature, shows that heterogeneity effects cannot be ignored.

Radioactive tracer methods have also been used to follow up adsorption processes. Wolf and Riehl,^{50, 152} for example, allowed graphite platelets of 0.1 mm. thickness to become exposed to Rn. After adsorption was complete, the system was brought in contact with a photographic plate. A marked blackening of the plate occurred at positions corresponding to the periphery of the platelets, while blackening in positions corresponding to the other surface portions was much less intense. Evidently, preferential adsorption at the edges and on the faces of the prism rather than on the basal planes is indicated. Other attempts to use this tracer method have been discussed by Crowell.³⁰

Calorimetric measurements provide additional experimental evidence. Results up to 1940 have been summarized in a review by Beebe⁸ and by Brunauer;¹⁹ numerous other publications have appeared in the literature since that time. In the majority of cases the differential or isosteric heats of adsorption, for physical adsorption of gases or vapors, fall from initial values of several k. cal./mole at low surface coverage to values of the order of 1 k. cal./mole at high coverage. It is unlikely that such a decrease can be explained solely in terms of lateral interactions. In the first place, much of the decrease occurs at such low coverage that lateral interactions are unlikely to be of major importance. Secondly, only repulsive lateral interactions can account for the progressive decrease in heats of adsorption. As explained in Section IV A (4), lateral repulsion can occur only in cases where gases with permanent multipoles are adsorbed on a surface; nevertheless, a fall in the heats of adsorption is usually observed in cases where electrostatic forces do not play a part. The above discussion does not, of course, rule out cases where adsorbent-adsorbate as well as lateral interactions act jointly (see Section VI).

Usually, the fall in heats of adsorption with increasing surface coverage is taken as evidence for the presence of a heterogeneous surface.

The change in apparent saturation characteristics of isotherms with temperature changes is regarded likewise as indicative of surface heterogeneity.^{19, 30, 55} This interpretation is discussed in detail by Crowell³⁰ and by Halsey.⁵⁵ Wilkins¹⁵¹ for example, determined adsorption isotherms for nitrogen, oxygen, and argon in the temperature range from 77.3 to 194.5° K. He noted a 10 to 100 fold change with temperature in the apparent quantity of gas required to cover the surface with a monolayer. A possible interpretation of these results is that the "weaker" sites, associated with small values of ϵ are not filled as readily at the higher temperatures as the "hot spots." Therefore, as the temperature is raised, only the latter remain covered. A more precise and quantitative reinterpretation of this statement will be given in Section V A (2). The saturation phenomena are only apparent, and at sufficiently high pressures the remainder of the surface will become covered. For further evidence concerning the existence of surface heterogeneity in this case the reader is referred to Halsey's paper.⁵⁵

Finally, deviations of isotherm data from the Langmuir equation are cited

as evidence for the existence of nonuniform surfaces. Such conclusions must be accepted with reserve, for reasons detailed in V B. However, if such deviations are very marked in the low pressure range it is at least reasonable to account for them in terms of surface heterogeneity. Likewise, the presence of steps in the adsorption isotherms^{8, 19, 76, 101} may be indicative of the existence of patches of sites associated with different ϵ values.⁵⁷ See Section VI A (2).

Occasionally, experiments are reported in the literature which seem to indicate that the effects of surface heterogeneity can be neglected. An example of this type is the meticulous work of Rhodin^{104, 105} concerning the sorption of N_2 on single crystal faces of copper and zinc. Electron microscope and diffraction studies showed that the surfaces exposed to the gas were essentially flat on an atomic scale; also, that their arrangement was essentially undistorted. The isotherms obtained on the 100, 110, and 111 faces under the same conditions differed considerably from each other. Likewise, the heats of adsorption were found to be different. A remarkable feature of these experiments is that there is no initial fall in the heat of adsorption with increasing surface coverage. This seems to confirm the energetic uniformity of the surface. A marked rise of the heats and a peak observed near monolayer coverage of the surface are attributable to the attractive lateral interaction potentials which become important upon progressive coverage of the surface.

On a polycrystalline surface, however, Rhodin observed a marked fall in the initial heats of adsorption with increasing surface coverage, and a much smaller hump in the heats near monolayer coverage. Evidently, the surface of the polycrystalline metal is energetically quite heterogeneous.

V. Distribution of Adsorption Energies Among Surface Sites

Having discussed theoretical and experimental evidence in support of the concept of surface heterogeneity, it seems desirable to develop methods to characterize its extent and its effect in gas adsorption. For a particular adsorbent-adsorbate system such a characterization is usually given in terms of a distribution function $F(\epsilon)$ whose differential $dF(\epsilon) = f(\epsilon) d\epsilon$ represents the fraction of surface sites associated with an adsorption energy in the range ϵ to $\epsilon + d\epsilon$. This distribution is generally not accessible to direct experimental observation, so indirect procedures must be used for its determination.

V A. Mathematical formalism based on the Langmuir concept. We turn first to the determination of the distribution function $dF(\epsilon)$ or of the distribution density $f(\epsilon) = dF(\epsilon)/d\epsilon$ from isotherm data, which specify the fraction of surface sites covered by adsorbate, θ , as a function of the equilibrium gas pressure, p , and of the temperature, T .

It will be assumed for the present that, on sites associated with a given value of ϵ , adsorption proceeds according to the well known Langmuir equation^{19, 31, 45, 48, 49, 51, 63, 72, 77, 78, 82, 117}

$$(5-1) \quad \theta_{\epsilon}(p, T) = \frac{p}{p + b(\epsilon, T)}$$

The theoretical treatment of cases based on a different set of assumptions

is discussed in Sections VI A and VI B. A critique of the present formalism is deferred to Section V B.

In EQUATION 5-1 the parameter b can be evaluated by use of statistical methods^{19, 48, 49, 51, 64, 65, 69}

$$(5-2) \quad b = \alpha(\epsilon, T) e^{-\epsilon/kT} \equiv \frac{(2\pi mkT)^{3/2}}{h^3} kT \frac{j_g(T)}{j_s(\epsilon, T)} e^{-\epsilon/kT}$$

where j_g and j_s represent the internal partition functions of molecules in the gas and adsorbed phases respectively.

On a heterogeneous surface, the total coverage is given by⁴⁹

$$(5-3) \quad \theta(p, T) = \int_{\epsilon_0}^{\epsilon_m} \theta_\epsilon dF(\epsilon) = \int_{\epsilon_0}^{\epsilon_m} \frac{dF(\epsilon)}{1 + \frac{\alpha}{p} e^{-\epsilon/kT}} = \int_{\epsilon_0}^{\epsilon_m} \frac{f(\epsilon) d\epsilon}{1 + \frac{\alpha}{p} e^{-\epsilon/kT}}$$

together with the normalization condition

$$(5-4) \quad \int_{\epsilon_0}^{\epsilon_m} dF(\epsilon) = \int_{\epsilon_0}^{\epsilon_m} f(\epsilon) d\epsilon = 1$$

EQUATIONS 5-3 and 5-4 have been given in two equivalent forms. Of these, the Stieltjes formulation in terms of the differential $dF(\epsilon)$, is the more general because it also covers those cases in which the variation of F with ϵ is discontinuous. This formulation can be illustrated as follows: suppose the surface consists of $n_1, n_2, \dots, n_m$ (m finite) sites with an associated set of discrete adsorption energies $\epsilon_1, \epsilon_2, \dots, \epsilon_m$; and let $n = \sum_{j=1}^m n_j$. A plot of $F(\epsilon)$ (the fraction of sites having adsorption energies in the range from 0 to ϵ) versus ϵ is shown in FIGURE 8. It consists of a series of m steps. Now, for $\epsilon \neq \epsilon_j$, *i.e.* in the region between discontinuities, $F(\epsilon)$ remains constant; hence $dF(\epsilon) = 0$, and no contribution to the integral $\int \theta_\epsilon dF(\epsilon)$ results. At each of the points of discontinuity $\epsilon = \epsilon_j$, however, $F(\epsilon)$ experiences a "jump," which corresponds to an increment $\Delta F(\epsilon_j) = n_j/n$ in $F(\epsilon)$. This increment contributes to the integral (5-3) an amount $\lim_{\delta \rightarrow 0} \int_{\epsilon_j - \delta}^{\epsilon_j + \delta} p dF(\epsilon) / [p + b(\epsilon, T)] = p \Delta F(\epsilon_j) / [p + b(\epsilon_j, T)] = p \frac{n_j}{n} / [p + b(\epsilon_j, T)]$. The total contribution is then the sum of all such increments at each of the points of discontinuity in $F(\epsilon)$. Thus,

$$(5-5) \quad \theta = \sum_{j=1}^n \frac{\Delta F(\epsilon_j)}{1 + b(\epsilon_j, T)/p}$$

Similarly, EQUATION 5-4 reduces to $\sum_{j=1}^m \Delta F(\epsilon_j) = 1$. For a rigorous treatment and discussion of the theory of Stieltjes integrals, the reader is referred to standard treatises.^{136, 149}

In EQUATIONS 5-3 and 5-4, the limits of integration were taken as ϵ_0 , and ϵ_m , the minimum and maximum adsorption energies encountered on the surface. The choice of limits is by no means unique. One could select any

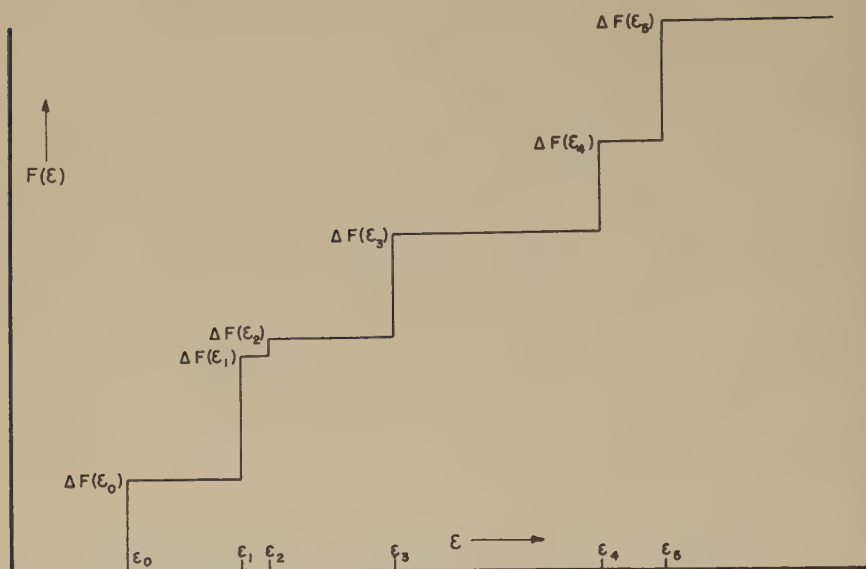


FIGURE 8. Distribution function $F(\epsilon)$ for adsorption energies among surface sites when distribution is discontinuous.

values ϵ_l , $-\infty \leq \epsilon_l \leq \epsilon_0$ and ϵ_u , $\epsilon_m \leq \epsilon_u \leq \infty$ as lower and upper limits respectively, provided one also requires that $F(\epsilon) = 0$ for $\epsilon_l \leq \epsilon < \epsilon_0$ and $F(\epsilon) = 1$ for $\epsilon_m < \epsilon \leq \epsilon_u$. Under these restrictions, the Stieltjes integral vanishes identically for $\epsilon < \epsilon_0$ or $\epsilon > \epsilon_m$, since $F(\epsilon)$ is constant in these ranges. The particular upper limit $F(\epsilon) = 1$ was chosen in conformity with the normalization requirement (5-4). Alternatively, one can replace the integration limits ϵ_0 , ϵ_m by ϵ_l , ϵ_u if one requires that $f(\epsilon) = 0$ outside the interval (ϵ_0, ϵ_m) .

From the definition of ϵ as the work required to remove an adsorbed atom or molecule from the surface to infinity, it follows that the physically meaningful values are restricted to the range $\epsilon \geq 0$. This point was emphasized by Hill^{68, 69} and again by Tompkins.¹³⁹ It is unfortunate that, in much of the theoretical work to be discussed later, $-\infty$ has been selected as the lower limit of integration without requiring simultaneously that $f(\epsilon)$ vanish over the range $-\infty \leq \epsilon < 0$. Retention of the limit $\epsilon_l = -\infty$ has lately again been favored as a mathematical convenience,⁶⁰ but the resulting mathematical simplifications obtain only if the requirement $f(\epsilon) = 0$ for $\epsilon < 0$ is also dropped. At best, elimination of this restriction can be justified in cases where $f(\epsilon)$ is small for $\epsilon < 0$.

It should be noted that the distribution functions in ϵ could be replaced by the distribution functions $G(b)$, $dG(b) = g(b) db$ which are defined by analogy to F , dF or f but describe distributions in the variable b rather than ϵ . EQUATIONS 5-3 and 5-4 then read

$$(5-6) \quad \theta(p, T) = \int_{b_0}^{b_m} \frac{dG(b)}{1 + b/p} = \int_{b_0}^{b_m} \frac{g(b) db}{1 + b/p}$$

$$(5-7) \quad \int_{b_0}^{b_m} dG(b) = \int_{b_0}^{b_m} g(b) db = 1$$

The use of distributions in b is favored in the Russian literature and presents a number of advantages; for, in order to solve EQUATION 5-3, the dependence of α on ϵ must be known. For simplification, it is usually assumed that α is independent of ϵ . On the basis of some theoretical estimates, Hill^{68, 69} has shown, however, that this assumption is untenable. A similar conclusion was reached by Drain and Morrison,^{40, 41} although these investigators are in disagreement with Hill as to the appropriate functional relationship. Even if the dependence of α on ϵ were properly specified, the solution of EQUATION 5-3 by the methods cited below would be very tedious.

Many of these troubles are avoided by solving EQUATION 5-6 for $G(b)$ or $g(b)$, since this problem is mathematically much simpler. If desired, one can later convert from $g(b)$ to $f(\epsilon)$ by use of EQUATION 5-2 as shown below, and it is only at that stage that the relation between α and ϵ is needed.

In order to prevent the occurrence of negative ϵ values, it is necessary to restrict b to the interval $0 \leq b \leq \alpha(0, T) \equiv \alpha_0$. In the Russian literature, the upper limit $b_m = \infty$ seems to have been used exclusively so far.

V A (1). Determination of distribution functions from isotherm data; exact methods. The functions F , f , G , or g may be determined by two methods. The first involves the use of trial and error procedures. One selects a number of trial functions for the distribution, subject to the normalization requirements, and carries out the indicated integration in EQUATIONS 5-3 or 5-6. The dependence of α on ϵ must also be given in advance for use in 5-3. The different functions $\theta(p, T)$ thus obtained are compared with experimental results and the function giving the closest fit is selected. The corresponding distribution function represents the correct variation of ϵ or b over the surface sites.

Langmuir,⁸² Pease,⁹⁷ and Foresti⁴⁷ had suggested the use of sums of the type (5-5) in cases where the data did not fit the Langmuir formula. Zeldowitsh¹⁵⁴ was the first to express surface coverage in terms of the integral EQUATION 5-6. Independently, Cremer and co-workers^{27, 28, 29} and later, Halsey and Taylor,⁵⁴ correlated $\theta(p, T)$ with $f(\epsilon)$ by the trial and error method. Zeldowitsh showed that use of the trial functions $g(b) = Ab^c$ ($0 \leq b \leq \infty$) in EQUATION 5-6 led to an isotherm first empirically proposed by Freundlich. Similarly, Cremer and Flügge,²⁷ also Halsey and Taylor,⁵⁴ verified that use of an exponential distribution law for $f(\epsilon)$ ($-\infty \leq \epsilon \leq \infty$), together with the assumption that α is independent of ϵ , leads to the Freundlich isotherm. Tompkins¹³⁹ and Hill⁶⁸ again rederived EQUATION 5-6 and have discussed its use in conjunction with the trial methods.

The second method for the determination of F , f , G , or g consists in solving EQUATIONS 5-3 or 5-6 directly by inversion of the integral. The first exact procedure for this purpose was introduced by Temkin and Levich¹³³ in 1946. It consists in rewriting EQUATION 5-6 as a Fourier integral and inverting the latter by standard mathematical procedures so as to solve for the unknown function under the integral sign. This method was further discussed by Todes

and Bondareva.¹³⁷ In 1948, Sips¹¹⁹ showed independently that EQUATION 5-3 could be reformulated as a Stieltjes transform of the distribution function and pointed out that standard methods are available¹⁴⁹ for the inversion of this type of integral. Sips later revised his theory¹²⁰ so as to exclude the range of negative ϵ values.

In either of the above methods it is necessary to represent the experimental isotherm data by an appropriate interpolation equation which is substituted for $\theta(p, T)$ on the left hand side of EQUATIONS 5-3 or 5-6. Upon bringing the integrals into the standard forms mentioned above one can solve for the desired distribution functions. This procedure will now be illustrated in detail by reviewing the article of Temkin and Levich.¹³³ A detailed summary of the method seems justified since their publication is in the Russian language and does not seem to be available in translation. Modifications of the original treatment have been introduced where necessary to exclude physically unacceptable b or ϵ values. Readers who do not wish to wade through all the mathematical details may go directly to EQUATION 5-21 where the important results and conclusions are summarized.

Since $0 \leq b \leq \alpha_0$ EQUATION 5-6 may be rewritten

$$(5-8) \quad \theta(p, T) = \int_0^{\alpha_0} \frac{pg(b) db}{p + b},$$

setting $g(b) = 0$ in the range $0 \leq b < b_0$ and $b_m < b \leq \alpha_0$, if necessary.

Use of substitutions

$$(5-9) \quad \xi = -\ln \left(\frac{\alpha_0}{b} - 1 \right) \quad \eta = -\ln \left(\frac{\alpha_0}{p} + 1 \right)$$

in (5-8) leads to

$$(5-10) \quad \theta \left(\frac{\alpha_0}{e^{-\eta} - 1}, T \right) \equiv L(\eta) = \int_{-\infty}^{\infty} \frac{\alpha_0 e^{-\xi} \cdot \sigma(\xi) \cdot d\xi}{(e^{-\xi} + e^{-\eta})(1 + e^{-\xi})} = \int_{-\infty}^{\infty} \frac{\varphi(\xi) d\xi}{1 + e^{\xi-\eta}}$$

where

$$(5-11) \quad \sigma(\xi) = g \left(\frac{\alpha_0}{1 + e^{-\xi}} \right) = g(b); \quad \varphi(\xi) = \frac{\alpha_0 \cdot \sigma(\xi)}{1 + e^{-\xi}} = b \cdot \sigma(\xi)$$

Differentiate (5-10) with respect to η , assuming that the differentiation under the integral sign is permissible,

$$(5-12) \quad \begin{aligned} L'(\eta) &= \int_{-\infty}^{\infty} \varphi(\xi) \cdot \frac{e^{\xi-\eta}}{(1 + e^{\xi-\eta})^2} d\xi \\ &= \int_{-\infty}^{\infty} \varphi(\xi) \cdot h'(|\eta - \xi|) d\xi \end{aligned}$$

It may be readily verified that the function $h' = e^{\xi-\eta}/(1 + e^{\xi-\eta})^2$ depends only on the absolute difference $|\xi - \eta|$.

We now write out the Fourier transform of $L'(\eta)$ which, by definition,¹³⁶ is

$$\begin{aligned}
 \mathbf{L}'(\nu) &= \int_{-\infty}^{\infty} e^{-2\pi i \nu \eta} L'(\eta) d\eta \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h'(|\eta - \xi|) \cdot \varphi(\xi) e^{-2\pi i \nu \eta} d\eta d\xi \\
 (5-13) \quad &\equiv \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{-2\pi i (\eta - \xi) \nu} h'(|\eta - \xi|) d\eta \right] e^{-2\pi i \nu \xi} \varphi(\xi) d\xi \\
 &= \int_{-\infty}^{\infty} H'(\nu) e^{-2\pi i \nu \xi} \varphi(\xi) d\xi = H'(\nu) \Phi(\nu)
 \end{aligned}$$

where H' and Φ are the Fourier transforms of h' and φ respectively.

To evaluate H' set $u = \xi - \eta$, $\zeta = e^u$ and $\gamma = 2\pi i \nu$. Then

$$(5-14) \quad H'(\nu) = \int_{-\infty}^{\infty} \frac{e^u}{(1 + e^u)^2} e^{2\pi i \nu u} du = \int_0^{\infty} \frac{\zeta^\gamma d\zeta}{(1 + \zeta)^2}$$

Integration by parts yields

$$(5-15) \quad H'(\nu) = \gamma \int_0^{\infty} \frac{\zeta^{\gamma-1}}{1 + \zeta} d\zeta = \frac{\gamma\pi}{\sin \gamma\pi} = \frac{2\pi^2 i \nu}{\sin 2\pi^2 i \nu} \quad (0 < \gamma < 1)$$

If $\gamma = 1$, $H'(\nu) = \int_0^{\infty} d\zeta / (1 + \zeta)^2 = 1$; if $\gamma = 0$, $H'(\nu) = \int_{-\infty}^{\infty} \frac{e^u du}{(1 + e^u)^2} = 1$ and if $-1 < \gamma < 0$, replacement of μ by $-u$ does not alter the final results. Using EQUATION 5-15 in 5-13, we find

$$(5-16) \quad \Phi(\nu) = \frac{\sin 2\pi^2 i \nu}{2\pi^2 i \nu} \mathbf{L}'(\nu)$$

According to well-known mathematical principles¹³⁶ the Fourier transform equation $\Phi(\nu) = \int_{-\infty}^{\infty} e^{-2\pi i \nu \xi} \varphi(\xi) d\xi$ may be inverted by writing $\varphi(\xi) = \int_{-\infty}^{\infty} \Phi(\nu) e^{2\pi i \nu \xi} d\nu$. Accordingly, we find from (5-16)

$$(5-17) \quad \varphi(\xi) = \int_{-\infty}^{\infty} [\sin 2\pi^2 i \nu / 2\pi^2 i \nu] \mathbf{L}'(\nu) e^{2\pi i \nu \xi} d\nu$$

Differentiation with respect to ξ yields

$$\begin{aligned}
 \varphi'(\xi) &= \int_{-\infty}^{\infty} \frac{\sin 2\pi^2 i \nu}{\pi} \mathbf{L}'(\nu) e^{2\pi i \nu \xi} d\nu \\
 (5-18) \quad &= \int_{-\infty}^{\infty} \frac{\mathbf{L}'(\nu)}{2\pi i} \{e^{2\pi i \nu (\xi + i\pi)} - e^{2\pi i \nu (\xi - i\pi)}\} d\nu
 \end{aligned}$$

By EQUATION 5-13, the two parts of the integral are proportional to the in-

verse Fourier transforms of $\mathbf{L}'$, i.e. so $L'(\xi + i\pi)$ and $L'(\xi - i\pi)$ respectively. We thus have finally

$$(5-19) \quad \varphi'(\xi) = \frac{1}{2\pi i} \{L'(\xi + i\pi) - L'(\xi - i\pi)\}$$

$$(5-20) \quad \varphi(\xi) = \frac{1}{2\pi i} \{L(\xi + i\pi) - L(\xi - i\pi)\}$$

The integration constant in (5-20) vanishes identically. For if we consider the value $\xi = -\infty$ which corresponds to $b = 0$, then $\theta_{b=0} = \int_{-\infty}^{\infty} g(b) db = 1$; accordingly, $\lim_{\xi \rightarrow \infty} L(\xi \pm i\pi) = 1$. But $\lim_{\xi \rightarrow \infty} \varphi(\xi)$ must vanish identically, since otherwise the integral (5-10) would diverge. Using these boundary conditions in EQUATION 5-20 one finds that the constant of integration is zero.

Substitution of EQUATIONS 5-11 and 5-9 into 5-20 yields the desired relation

$$(5-21) \quad g(b) = \frac{1}{2\pi i b} \left\{ L \left[-\ln \left(\frac{\alpha_0}{b} - 1 \right) + \pi i \right] - L \left[-\ln \left(\frac{\alpha_0}{b} - 1 \right) - \pi i \right] \right\}$$

EQUATION 5-21 is a solution of the integral EQUATION 5-6. The function $L(\eta)$ is obtained from the interpolation equation $\theta(p, T)$ upon replacing p with $\alpha_0/(e^{-\eta} - 1)$ (See EQUATION 5-10). Next, one replaces the variable η , wherever it occurs in the function L , by the quantity $-\ln \left(\frac{\alpha_0}{b} - 1 \right) \pm \pi i$; one can then solve for $g(b)$ in EQUATION 5-21.

One may pass from g to f by noting that a set of ϵ values in the range ϵ_i to $\epsilon_i + d\epsilon_i$ is in 1:1 correspondence with a set of b values in the range b_i to $b_i + db_i$; where db_i and $d\epsilon_i$ are related by use of EQUATION 5-2. Hence the number of sites associated with this range of ϵ or b values is the same, or $dF(\epsilon) = dG(b)$. Leaving the dependence of α on ϵ unspecified for the moment, one finds from EQUATION 5-2

$$(5-22) \quad \begin{aligned} f &= -g \left(\frac{\partial b}{\partial \epsilon} \right)_T = -g\alpha(\epsilon, T) \cdot e^{-\epsilon/kT} \cdot \left(-\frac{1}{kT} + \frac{\partial \ln \alpha}{\partial \epsilon} \right) \\ &= -gb \left(-\frac{1}{kT} + \frac{\partial \ln \alpha}{\partial \epsilon} \right) \end{aligned}$$

The minus sign has been introduced into (5-22) to take account of the fact that an increase in ϵ by $d\epsilon$ is matched by a decrease in b by db .

Using relation (5-22) one can then write

$$(5-23) \quad \begin{aligned} f(\epsilon) &= \frac{1}{2\pi i} \left(-\frac{1}{kT} + \frac{\partial \ln \alpha}{\partial \epsilon} \right) \left\{ L \left[-\ln \left(\frac{\alpha_0}{\alpha} e^{\epsilon/kT} - 1 \right) - \pi i \right] \right. \\ &\quad \left. - L \left[-\ln \left(\frac{\alpha_0}{\alpha} e^{\epsilon/kT} - 1 \right) + \pi i \right] \right\} \end{aligned}$$

It is only now that the dependence of α on ϵ must be specified to obtain an

explicit form of $f(\epsilon)$. If α is assumed independent of ϵ , $\alpha = \alpha_0$, $b = \alpha_0 e^{-\epsilon/kT}$ and EQUATION 5-23 simplifies to

$$(5-24) \quad f(\epsilon) = \frac{1}{2\pi i kT} \{L[-\ln(e^{\epsilon/kT} - 1) - \pi i] - L[-\ln(e^{\epsilon/kT} - 1) + \pi i]\}$$

which is in a form equivalent to the relation derived by Sips.¹²⁰

The use of this theory may be illustrated by a simple example. Suppose the experimental data can be represented by an empirical relation of the form

$$(5-25) \quad \theta(p, T) = \left(\frac{p}{p + \alpha_0(T)} \right)^c \quad (0 < c < 1)$$

This relation, first proposed by Sips,¹²⁰ approaches the Freundlich equation at sufficiently low pressures. Use of the definitions (5-9) leads to

$$(5-26) \quad \theta\left(\frac{\alpha_0}{e^{-\eta} - 1}, T\right) \equiv L(\eta) = e^{c\eta}$$

When this relation is used in EQUATION 5-21 the desired distribution density function reads

$$(5-27) \quad g(b) = \frac{\sin \pi c}{\pi} \left(\frac{1}{b} \right) \left(\frac{\alpha_0}{b} - 1 \right)^{-c}$$

Converting to $f(\epsilon)$ we find

$$(5-28) \quad f = \frac{\sin \pi c}{\pi} \left[\frac{1}{kT} - \frac{\partial \ln \alpha}{\partial \epsilon} \right] \cdot \left(\frac{\alpha_0}{\alpha} e^{\epsilon/kT} - 1 \right)^{-c}$$

If one introduces the restriction $\alpha = \alpha_0$, independent of ϵ , one finds

$$(5-29) \quad f(\epsilon) = \frac{\sin \pi c}{\pi kT} (e^{\epsilon/kT} - 1)^{-c}$$

in agreement with the result by Sips. Thus, (5-29) is the distribution which, under the assumed independence of α from ϵ , corresponds to the generalized Freundlich relation (5-25). Specification of other relations between α and ϵ will lead to different distribution densities.

The equivalent inversion procedure used by Sips is explained in detail in his papers^{119, 120} and need not be discussed further.

If F or G is known to be a step function [i.e., if the energy spectrum consists of a discrete set of values, $\epsilon_j (j = 1, 2, 3, \dots, n)$], the distribution may be determined using a method discussed by Todes and Bondareva.¹³⁷ One selects m experimental points $\theta_1, \dots, \theta_\alpha, \dots, \theta_m$, corresponding to the pressure values $p_1, \dots, p_\alpha, \dots, p_m$. The adsorption isotherm (5-5) then reads

$$(5-30) \quad \theta_\alpha \equiv \theta(p_\alpha) = \sum_{k=1}^m \frac{p_\alpha}{p_\alpha + b_k} \Delta G_k \quad \alpha = 1, \dots, m$$

subject to the condition

$$(5-31) \quad \sum_{k=1}^m \Delta G_k = 1$$

EQUATIONS 5-30 and 5-31 represent a system of nonlinear algebraic equations which, in principle at least, can be solved for the unknowns ΔG_k .

V A (2). *Determination of distribution functions from isotherm data; approximation methods.* For reasons discussed in Section V B, approximation procedures frequently serve almost as well as the exact methods in correlating distribution densities f or g with a given set of isotherm data. Such methods also illustrate other principles of interest and are discussed briefly below.

Temkin and Levich¹³³ pointed out that EQUATION 5-20 could be expanded in a Taylor's series about ξ ; formally this leads to the relation:

$$(5-32) \quad \varphi(\xi) \approx L'(\xi) - \frac{\pi^2}{3!} L'''(\xi) + \frac{\pi^4}{5!} L^{(5)}(\xi) - \dots$$

If only the first term is retained in this expansion one obtains the formulation proposed independently by Roginskii.¹⁰⁹ Roginskii called attention to the fact that a plot of θ versus ϵ , according to EQUATION 5-1, yields a series of sigmoidal curves at a series of pressures. These curves are reproduced in FIGURE 9. Inspection of this graph shows that, at $p = 10^{-3}\alpha_0$, for example,

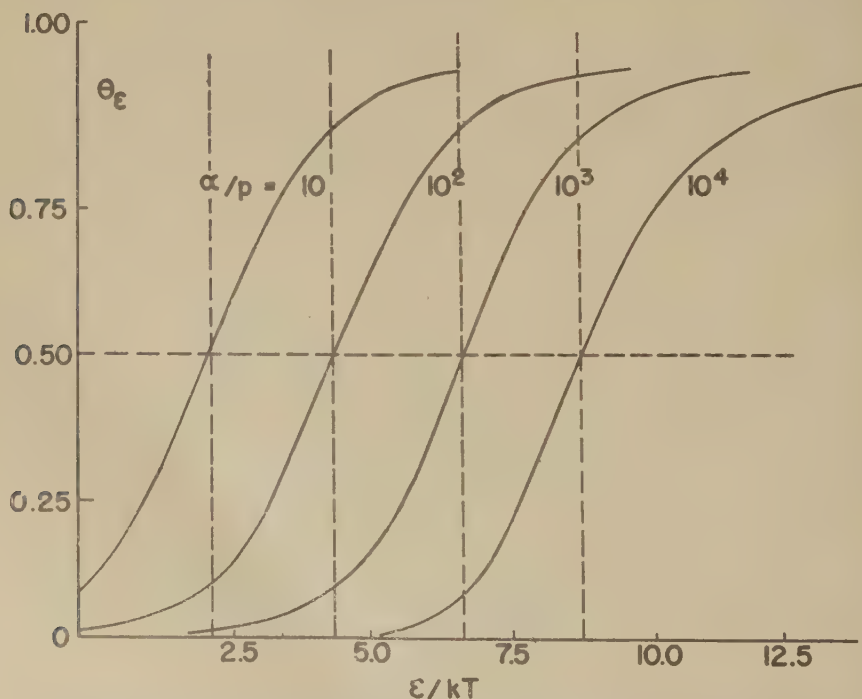


FIGURE 9. A plot of θ versus ϵ according to EQUATION 5-1, at different pressures. In Roginskii's method the curves are replaced by dotted lines shown on the diagrams.

the degree of occupation of sites associated with $\epsilon = 4.7, 6.1, 6.9, 7.8$, and 9.1 kT is $\theta_\epsilon = 0.1, 0.3, 0.5, 0.7$, and 0.9 respectively. Thus, all sites for which $\epsilon < 4.7$ kT or for which $\epsilon > 9.1$ kT (a spread of 4.4 kT units), are effectively completely empty or completely filled at that particular pressure. Quite generally, it is only the sites with associated ϵ values in a range of $\Delta\epsilon/kT \approx 4.4$ whose occupancy is appreciably altered as the pressure is raised. Thus, if the total spread of ϵ values over the entire surface is much larger than 4.4 kT , it is permissible to replace the curves of FIGURE 9 by the dotted vertical lines at $\epsilon = \epsilon_i$ which intersect the original curves at $\theta_i = 1/2$. In this approximation, one thus sets $\theta_i = 0$ for $\epsilon < \epsilon_i$ and $\theta_i = 1$ for $\epsilon > \epsilon_i$. Thus from (5-3)

$$(5-33) \quad \theta = \int_{\epsilon_i}^{\infty} f(\epsilon) d\epsilon$$

The integral will converge if $\lim_{\epsilon \rightarrow \infty} f(\epsilon) = 0$. Under this condition, one can invert the integral to find

$$(5-34) \quad f(\epsilon) = -\left(\frac{\partial \theta}{\partial \epsilon}\right)_{\epsilon=\epsilon_i}$$

Now in order that $\theta_{\epsilon_i} = 1/2$, one must set $p_i = b_i = \alpha_0 e^{-\epsilon_i/kT}$ in EQUATION 5-1. One now obtains from (5-34)

$$(5-35) \quad f(\epsilon) = -\left(\frac{\partial \theta}{\partial p} \frac{\partial p}{\partial \epsilon}\right)_{\epsilon=\epsilon_i} = \frac{1}{kT} \left(p \frac{\partial \theta}{\partial p}\right)_{p=\alpha_0 e^{-\epsilon_i/kT}}$$

This type of approximation has also been used by Halsey⁵⁶ and by Schilling.¹¹⁴ The approximation becomes clearly inapplicable at pressures for which sites with ϵ values near 0 begin to be filled appreciably. EQUATION 5-35 can be handled much more readily than relations 5-21 or 5-22. For comparison it might be noted that use of the Freundlich isotherm $\theta = k p^c$ in EQUATION 5-35 corresponds to a distribution density f which is proportional to $\exp(-c\epsilon/kT)$. On the other hand, use of the generalized Freundlich relation 5-25 in the exact theory leads to an f which is proportional to $[\exp(\epsilon/kT) - 1]^{-c}$ (see EQUATION 5-29).

Roginskii and Todes¹¹¹ suggest an approximation in which the quantity $h'(|\eta - \xi|)$ of EQUATION 5-12 is replaced by the step function $\bar{h}' = 1/4$ for $\eta - 2 \leq \xi \leq \eta + 2$ and $\bar{h}' = 0$ for all other ξ . EQUATION 5-12 now reads

$$(5-36) \quad 4L'(\eta) = \int_{\eta-2}^{\eta+2} \varphi(\xi) d\xi$$

and this relation is readily inverted:

$$(5-37) \quad \varphi(\eta + 2) - \varphi(\eta - 2) \Big|_{\eta=\xi} = 4L''(\eta) \Big|_{\eta=\xi}$$

Owing to the existence of a second derivative on the right and since $\varphi(\eta)$ is not evaluated directly, EQUATION 5-37 seems to have little practical value.

Still another approximation was introduced by Zeldowitch,¹⁵¹ who replaced EQUATION 5-1 with two relations

$$(5-38) \quad \begin{aligned} \theta_{\epsilon_i} &= p/b & \text{for } 0 < p < b \\ \theta_{\epsilon_i} &= 1 & \text{for } p > b \end{aligned}$$

In effect, the approximation consists of extrapolating the initial slope of the isotherm for each set of sites of energy ϵ_i to the value $\theta_{\epsilon_i} = 1$. Then it can be shown that corresponding to

$$(5-39) \quad \theta(p) = \int_0^p \frac{p g(b) db}{b} + \int_p^{\alpha_0} g(b) db$$

the distribution density reads

$$(5-40) \quad g = -p \theta''(p) \big|_{p=b}$$

Zeldowitch¹⁵⁴ demonstrated that use of the Freundlich isotherm in EQUATION 5-40 leads to an exponentially decreasing distribution density g .

Finally, Todes and Bondareva¹³⁷ suggested the use of a very elegant procedure which involves expansion of experimental isotherm data in the high pressure range $p \rightarrow \infty$, $\theta \rightarrow 1$ in terms of the power series

$$(5-41) \quad \theta(p) = \sum_{j=0}^{\infty} K_j / p^j$$

It is possible to show that the coefficients K_j are given by the relations

$$(5-42) \quad K_j = \int_0^{\infty} b^j dG(b)$$

and that (5-42) can be inverted. The procedure thus consists of finding the K_j by fitting the experimental data to EQUATION 5-41, and then solving the integral EQUATION 5-42 for $dG(b)$. Details of this method are omitted, since its application is about as complicated as the exact methods. Furthermore, use of data in the high pressure range leads to many other difficulties discussed in Section V B (1).

V B. Critique of the inversion procedure. We discuss next the limitations inherent in the formalistic inversion procedures described above.

V B (1). Limitations attributable to assumptions implicit in the Langmuir theory. The theory of Section V A is based on the hypothesis that all sites of equal energy are occupied in accordance with the Langmuir law (5-1). The theory of Section V A is thus subject to the following assumptions: (a) adsorption occurs on a fixed number of sites without perturbing the adsorbent; (b) adsorption is limited to a monolayer; (c) no lateral interactions between adsorbed molecules occur; and (d) each adsorbed particle is rigidly bound to its adsorption site, no lateral motion occurring on the surface.

The simultaneous applicability of all these assumptions is open to doubt for most adsorbate-adsorbent systems. It is generally assumed that (a) represents a reasonable approximation.⁶⁹ However, Cook, Pack, and Oblad²⁴ have postulated that atoms in the surface layer are arranged in disorderly, strained patterns. The arrangement is thought to be such that the surface atoms with free valencies are given a chance to establish multiple linkages with each other. This surface strain is relieved during adsorption by the formation of covalent or ionic linkages of the adsorbate with the surface atoms.

The energy released in the ordering process is much larger than that accompanying physical adsorption on the remaining, ordered sites, and this effect is postulated to be the physical basis for the existence of nonuniform surfaces.

As was shown in Section IV A, any unevenness in the topography of the surface gives rise to a spread in ϵ values of the order of one magnitude. On a sufficiently porous surface, an appreciable fraction of sites is associated with large ϵ values. It is thus not essential to invoke the special model of these authors in order to explain the heterogeneity effects.

The postulated irregularity in the arrangement of surface atoms should be detectable experimentally. Cases are known,^{35, 36} however, in which the ordered arrangement of the bulk of the crystal seems essentially preserved right up to the surface layer, although the spacings may be distorted. Furthermore, in a series of publications (see for example¹¹⁶) Selwood has demonstrated that small quantities of a solid, dispersed in agglomerates over a support, tend to assume the crystal structure of the underlying support in preference to their own. It is difficult to explain this phenomenon of "valency induction" on the assumption of a disordered arrangement of surface atoms.

The hypothesis of Cook, Pack, and Oblad does point to an important fact: a clearcut distinction between chemisorption or physical adsorption processes frequently cannot be made. These two types of adsorption may occur side by side. Evidence for such adsorption phenomena is reviewed by de Boer.³⁵ Alternatively, the adsorbent may swell or a rearrangement of the surface atoms during adsorption may take place.¹⁹

Concerning assumptions (a) and (d), the "site" concept has meaning only if the adsorbate is essentially localized near potential minima of the surface. If the effective area of the adsorbate is larger than that of a "site," adsorption may stop before the surface is completely covered, or else a rearrangement of adsorbed atoms may occur to accommodate further atoms on the surface. In neither case will the calculated F or f function represent a true distribution function over all sites. If appreciable lateral motion sets in, the adsorbate molecules will experience continuous variations in ϵ during their travel. In such cases, the use of the Langmuir EQUATION 5-1 is manifestly incorrect.

A number of criteria are available to ascertain if, in a given case, adsorption is localized or mobile. The field emission microscope may be used for direct visual observation in some cases.^{2, 7} Other special methods are reviewed by Brunauer¹⁹ (chapter 13).

Where direct experimental methods are not available, indirect procedures may be used. This requires (a) an estimate of barrier heights hindering lateral translation; and (b) a comparison to the thermal energy kT :

(a) For the estimates of barrier heights a rule of thumb is often useful: at a given site, the adsorbate is in direct contact with Z nearest neighbors; motion to an adjacent site requires a temporary rupture of one or more direct contacts. If n contacts remain during transition, the height of the energy barrier ΔW_0 is roughly $\frac{Z-n}{Z} \epsilon$. (See also de Boer.³⁶) This estimate is based

on the approximation that the entire interaction energy derives from the direct contracts. Improved values are obtained by use of the contour diagram

method discussed in Section II D. Unfortunately, it is difficult to compute the magnitude of barriers on nonuniform surfaces, since the geographic distribution of the surface features must be specified to construct the contour diagram.

(b) There then remains the task of deciding at what values of kT relative to ΔW_0 mobility sets in. It is customary to assume that a transition between localized and mobile adsorption starts in at $kT = \Delta W_0/10$.^{37, 139} This estimate was confirmed by Hill,⁶⁵ who calculated the heat capacity C_s for the two lateral degrees of freedom of an atom adsorbed on a smooth surface. His plot of C_s versus T passed through a maximum in the region $T \approx \Delta W_0/10k$. Thus, if ΔW_0 is taken as 500 cal./mole the transition interval should commence at liquid hydrogen temperatures.

Experimentally, one should be able to define the transition region by measuring C_s as a function of T . Unfortunately, where such measurements have been carried out^{40, 41} C_s was found to rise with T but did not pass through a maximum. Here, the presence of "configurational effects" discussed by Drain and Morrison⁴¹ complicated the interpretation of the experimental data.

Fortunately, there is another method that can be used as a criterion for lateral mobility. As described by Kemball,⁷⁹ de Boer and Kruyer,³⁷ and by de Boer,³⁶ the procedure consists in determining the entropy of the adsorbed gas molecules. These values can be correlated with the number and types of degrees of freedom of the adsorbate molecules on the surface, from which conclusions as to lateral mobility can be drawn. Such calculations have been carried out by now for a large number of adsorbent-adsorbate systems. The results indicate that truly localized adsorption (gas molecules oscillate laterally about the equilibrium position but do not pass over potential barriers to adjacent sites) is usually encountered only in the temperature range of liquid hydrogen or below.

The inversion procedures discussed in Section V A should therefore be applied only to isotherms obtained at low temperatures. Numerical calculations by Honig and Reyerson⁷⁰ or by Wiig and Smith¹⁵⁰ must be accepted with reserve, since under their experimental conditions assumption (d) was no longer strictly applicable. It is shown in Section VI B, however, that the formalism of Section V A (1) is not greatly altered if mobility is taken into account.

Assumption (b) represents another very serious restriction, since the majority of physical adsorption studies at low temperature involve multilayer formation. It is, of course, possible to avoid this difficulty by selecting data from the low pressure region of the isotherm. These data must lie well below the point where monolayer coverage is completed, since second layer formation often sets in before all of the bare surface is covered. Restriction of the inversion procedure to the low pressure region of the isotherm introduces new difficulties, however, which will be discussed in section V B (2). To eliminate the impasse, one would have to subtract the multilayer contribution from the total isotherm and to construct $\theta(p, T)$ from the corrected monolayer isotherm. This alternative is hardly practicable, however.

A number of attempts have been made to lift restriction (b). These attempts have been based on the use of the Brunauer, Emmett, and Teller (BET) equation, which is an extension of the Langmuir concept to multilayer forma-

tion on a uniform surface. This equation has the form^{18, 19, 22, 36, 65}

$$(5-43) \quad \theta(x, T) = \frac{cx}{(1-x)(1-x+cx)}$$

where⁸⁴

$$(5-44) \quad c = \frac{j_1}{j_L} e^{(\epsilon_1 - \epsilon_L)/kT} \quad x = p/p_0$$

p_0 being the vapor pressure of the adsorbate at the temperature of the sample, and the subscripts 1 and L referring to the adsorbate in the first layer and in the liquid state respectively. One of the basic assumptions in the derivation of EQUATION 5-43 is that the second and higher layers have the properties of a liquid. This equation has been generalized by de Boer.³⁶

To apply (5-43) to heterogeneous surfaces McMillan,⁹⁴ also Walker and Zettlemoyer,¹⁴⁵ postulated a dual surface model with widely differing adsorption energies for each segment. The isotherm consists of two terms (5-43) weighted according to the relative abundance of each type of sites. Honig⁷¹ introduced another simple extension by writing out the equation $\theta(x, T) =$

$\int_{\epsilon_0}^{\epsilon_m} \theta_f(\epsilon) d\epsilon$ where θ_ϵ is given by (5-43). Following the procedure of Sips,¹²⁰

he showed that this integral would be inverted and solved for $f(\epsilon)$. The above methods can be criticized on the grounds that the original BET theory is itself quite inadequate, since it neglects precisely the lateral interactions which are required for the formation of an adsorbed liquid in the higher layers. Moreover, in Honig's treatment, the variation in f over sites in the second and higher layers is ignored. This approximation is not unreasonable since ϵ falls off approximately as z^3 (see EQUATION 2-17). The spread in ϵ values for second and higher layer sites is therefore much less than that for adsorption on the bare surface. Nevertheless, the transmission of surface heterogeneity effects through several adsorbed layers cannot be ignored in any adequate theoretical treatment. Further discussion of this topic is deferred to Section VI A (2), where more refined theories of multilayer sorption are discussed.

Assumption (c) again very seriously restricts the usefulness of the inversion theory. Adsorbate molecules in sufficient proximity experience van der Waals' forces of attraction. Additional interaction forces arise if the adsorbate possesses permanent multipoles. If attractive interactions predominate, two dimensional condensation below the critical temperature may set in (see Section VI). Any theory based on the use of the Langmuir equation is obviously incapable of describing this phenomenon. For this reason, use of the inversion theories of Section V A should be confined to cases where the lateral interaction energies are at least small compared to the adsorbent-adsorbate binding energies. Alternatively, use of the inversion should be restricted to the range of low surface coverage, where the distance between adsorbate molecules is large.

Theories which attempt to treat surface heterogeneity and lateral interaction effects simultaneously are presented in Section VI.

Finally, it must be recalled that the distribution functions calculated by the

methods of Section V A may vary for different adsorbates [see Sections III D (1) and IV A (4)]. This is particularly noticeable when van der Waals and electrostatic interactions predominate in two different cases. Again, gases which are small enough to fit into crevices inaccessible to a larger species may yield characteristic distribution functions not obtained with the larger molecules. In other instances, certain forces contributing to adsorption of one species may be lacking in others. An example of this sort is found in comparing the adsorption of N_2 with that of O_2 or A. The heats of adsorption of N_2 on solids are far larger than that of O_2 or A.⁷⁰ This has been explained by Drain⁴² in terms of quadrupole interactions which contribute noticeably to the N_2 but not to O_2 or A interactions with the solid. A given distribution function thus gives the variation of ϵ over a solid surface with reference to a particular adsorbate.

V B (2). *Mathematical limitations of the method.* Several other factors restrict the usefulness of the inversion procedure. These factors are discussed extensively in an article by Roginskii and Todes.¹¹¹

(a) The isotherm data themselves cannot be used to differentiate between a strictly homogeneous and a markedly heterogeneous surface. Vol'kenshtein^{143, 138} and Everett⁴⁶ have developed methods which relate a given isotherm with a particular law of lateral interaction or a particular equation of state for the adsorbate on a smooth surface. Thus, the same adsorption isotherm which is used in Section V A to determine the extent of the surface heterogeneity under the assumption (d) is used by the above authors to determine the extent of the lateral interactions on a uniform surface. Obviously, the inversion procedures must not be applied unless it is known from other evidence that the surface is nonuniform.

(b) It must be recognized that use of the inversion theory puts very exacting requirements on the data used in the empirical interpolation equation $\theta(p, T)$. Since integration processes always tend to "smooth out" very severe fluctuations of the integrand it is evident that the function $\theta(p, T)$ is very insensitive to large fluctuations in $f(\epsilon)$. Conversely, $f(\epsilon)$ cannot be determined unless θ is known very accurately.

One is thus faced with two requirements: first, the selected relation $\theta(p, T)$ must represent the best possible empirical fit to the experimental data; it will not do to select a relation which passes only through most of the experimental points. Unless the data are very accurately described by an interpolation relation, there is no assurance that the calculated and real distributions will agree even remotely. Obviously, the experimental work must be done with a high order of precision to permit close curve fitting.

Second, it is necessary to obtain data over a wide range of experimental conditions. As Todes and Bondareva¹³⁷ point out, if such data are not secured, several interpolation equations can generally be made to fit the data equally well in the restricted range. These authors conclude, for example, that Dubinin and Chrenova's adsorption isotherms for ethyl alcohol on various solids⁴⁴ can be represented equally well by the Freundlich, Langmuir, or Schmidt isotherms over a sixfold change of vapor pressures. Obviously, an

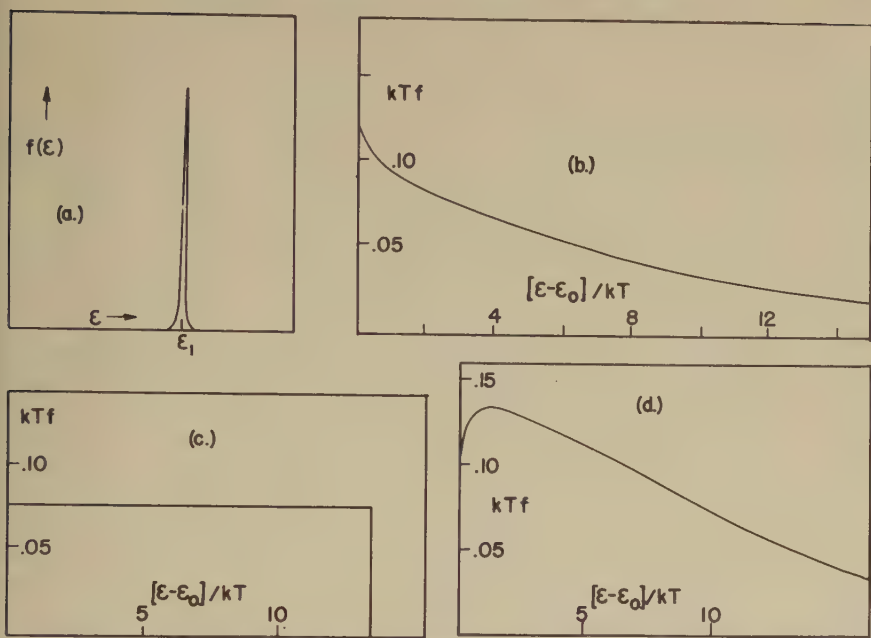


FIGURE 10. Four types of distribution densities $f(\epsilon)$ used as trial functions in the integral equations to reproduce experimental data shown in FIGURE 11.

assignment of suitable distribution functions is not possible on the basis of such limited data.

Similar conclusions were reached by Honig⁷¹ who attempted to determine $f(\epsilon)$ from nitrogen adsorption isotherms on rutile at 77° K. Several radically different trial functions (see FIGURE 10) fitted the isotherm data equally well over the limited range of surface coverage $0.4 \leq \theta \leq 1.4$, as is shown in FIGURE 11. The best that could be done was to divide the distribution functions into sets compatible or incompatible with the data. Among the latter (see FIGURE 11, full circles) were the function $f(\epsilon) = \delta(\epsilon - \epsilon_1)$ (Dirac delta function, corresponding to the assumption of a uniform surface) and all functions which remain bounded away from zero as $\epsilon \rightarrow \infty$. This latter class of functions renders the integrals divergent.

These considerations emphasize the rigid requirements that must be met by isotherm data which are to be used in inversion procedures. Unless these conditions are met, the approximation procedures of Section V A (2) serve just as well as the exact methods to obtain a qualitative trend for f with ϵ .

In this connection it might be of interest to cite unpublished work by Honig and Rosenbloom. It is possible to select two distribution densities f_a and f_b for use in EQUATION 5-3 such that the corresponding integrals $\theta_a(p, T)$ and $\theta_b(p, T)$ are extremal. Thus, for any other $f \neq f_a, f_b$ the corresponding isotherm $\theta(p, T)$ will lie within the region on a θ versus p diagram bounded by θ_a and θ_b . One obtains, therefore, an estimate on the maximum variation in

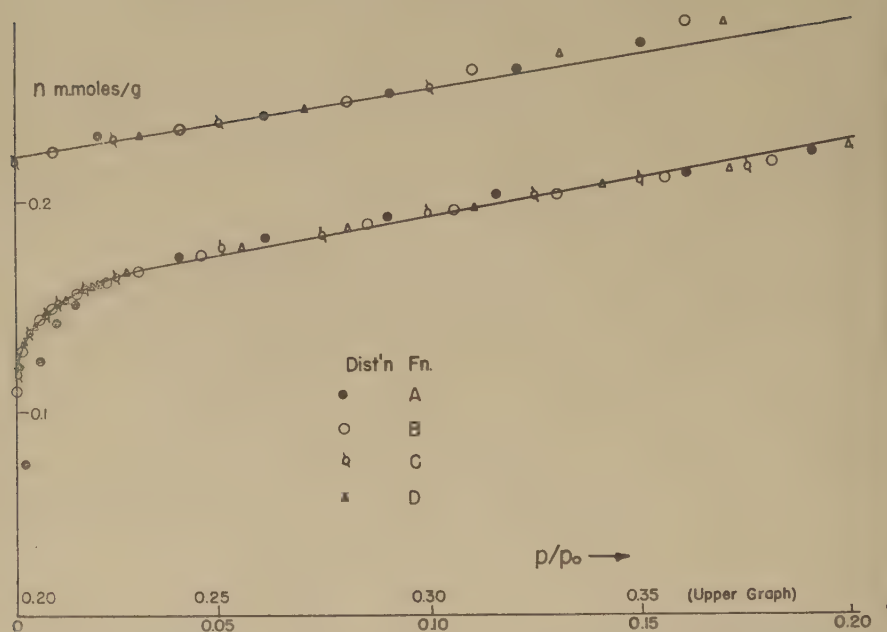


FIGURE 11. Adsorption isotherm for N_2 on rutile (TiO_2) at $77.7^\circ K$. Solid line interpolation isotherm representing the experimental data. Circles are points on isotherm computed by use of the distribution density functions (a) to (d) of FIGURE 9. Note deviation of solid circles [distribution (a), representing a Dirac delta function that corresponds to a homogeneous surface] from isotherm data at low pressure.

$\theta(p, T)$ which can be expected as f is allowed to assume all conceivable relationships that leave the integral (5-3) convergent. Conversely, such a procedure may be used as a check on the applicability of assumptions (a) to (d): after constructing the functions θ_a and θ_b from a small number of experimental points, one observes whether the remaining points fall within the limits just computed. If such is not the case, some of the assumptions (a) to (d) are not compatible with the experimental data.

The collection of data in the low and high pressure regions also presents the following advantage: as was shown in Section V A (2), (FIGURE 9) only sites whose associated energy falls in the narrow range of $4.4 kT$ units experience a change in occupancy as the pressure is altered slightly. Hence, it is necessary to obtain data at very low and very high pressures in order to determine the relative abundance of "hot" and "cold spots." Such information cannot be obtained from an intermediate pressure range in which the occupancy of sites at either end of the energy spectrum remains essentially unaltered.

(c) A number of purely mathematical restrictions must also be kept in mind. These have been reviewed by Sips¹²⁰ and by Todes and Bondareva.¹³⁷ (i) $\lim_{\epsilon \rightarrow \infty} f(\epsilon) = \lim_{b \rightarrow 0} g(b) = 0$, otherwise the integral (5-3) or (5-6) diverges; alternatively $F(0) = 0$ and $\lim_{\epsilon \rightarrow \infty} F(\epsilon) = 1$. (ii) At very low pressures, $p \ll b(\epsilon)$ for all ϵ , the isotherm must have the form $\theta(p, T) \approx p \int_0^\infty dF(\epsilon) / \alpha e^{-\epsilon/kT} = p \overline{b(1/b)}$. Thus b must be chosen so that the average

value of $1/b$ is given by the slope of the initial straight line portion of the isotherm. (iii) At high pressure, $p \gg b(\epsilon)$ for all ϵ , EQUATION 5-6 simplifies to $\theta \approx 1 - \frac{1}{p} \int_0^\infty \alpha e^{-\epsilon/kT} dF(\epsilon) = 1 - \bar{b}/p$; b must be so chosen that its average satisfies the asymptotic high pressure relation $\theta = 1 - \bar{b}/p$. (iv) By direct differentiation of (5-6) one sees that $\theta(p, T)$ and all derivatives are alternately monotonically increasing or decreasing functions of p . Interpolation formulae must satisfy this condition. (v) Considering p as a complex number whose real part has physical significance, it can be shown that $\text{Im } \theta(p) > 0$ for $\text{Im } |p| > 0$.¹⁴⁹ (vi) Physical as well as mathematical considerations require that $\theta(p, T) \geq 0$ for $0 \leq p \leq \infty$. (vii) It may be proved^{120, 149} that, if the Sips inversion procedure is employed, the quantity $L(\eta) = \theta(\alpha_0/(e^{-\eta} - 1), T)$ must be an analytic function and single valued in the complex plane except for a cut along the negative real η axis from 0 to infinity.

All the above requirements drastically limit the interpolation equations, which may be used to represent experimental data for use in inversions. Todes and Bondareva¹³⁷ point out that the Freundlich relation, the logarithmic isotherm $\theta = (C_0 + C_1 \ln p)^m$, the Bradley-Palmer-de Boer and Zwikker relation $\theta = C_1 \ln \ln p + C_2$, the Harkins and Jura isotherm $\theta = (k_1 - k_2 \ln p)^{1/2}$ and the Schmidt isotherm do not satisfy all of the above requirements.

(d) Finally, one should note that the distribution function f or F calculated by the above inversion procedures will usually contain the temperature T as a parameter. This is illustrated by the derivation of EQUATION 5-29 from use of (5-25) in (5-24). It appears difficult to furnish a physical justification for such a dependence. For, unless deliberate steps are taken to alter the surface by high temperature treatment, the distribution of sites should not be affected by the adsorption temperature. In unpublished work by Honig and Edward L. Hill the view is taken that f should deliberately be made independent of T . This requirement further restricts the possible interpolation relations which may be used to represent the isotherm data. It also enables one, in principle, to determine both $f(\epsilon)$ and the dependence of α on ϵ if the pressure and temperature dependence $\theta(p, T)$ is known.

VI. Gas Adsorption on Heterogeneous Surfaces in the Presence of Lateral Interactions

The theory of Section V can be considerably refined by taking account of lateral interactions among adsorbed molecules. Unfortunately, a rigorous theory of this process becomes rather untractable. It is usually necessary to introduce simplifying assumptions or special models to obtain explicit solutions.

VI A. Fixed lateral interactions. We shall examine first a fairly crude model: assumption (c) of Section V B (1) will be replaced by the assumptions (e) that molecules on adjacent surface sites experience a fixed lateral interaction potential w , and (f) that $w = 0$ for molecules on nonadjacent surface sites.

The theory for localized, monolayer sorption on uniform surfaces under assumptions (e) and (f) has been discussed by Roberts¹⁰⁸, Miller⁹⁵ and by Fowler and Guggenheim.⁴⁹ The latter treatment has been extended by Hill and by

Halsey to adsorption on heterogeneous surfaces. Some of the important conclusions of the Fowler and Guggenheim theory will therefore be discussed briefly.

VI A (1). *Adsorption on uniform surfaces.* Fowler and Guggenheim⁴⁹ used two different statistical approximations, the Bragg-Williams (BW) and the quasi-chemical (QC) approximations. These are reviewed briefly by Hill,^{68, 69} and are treated more extensively elsewhere.^{49, 107} Using these approximations, Fowler and Guggenheim derived the relations*

$$(6-1) \quad p(\theta) = b \frac{\theta}{1-\theta} e^{zw\theta/kT} \quad (B-W) \quad (w < 0)$$

$$(6-2) \quad p(\theta) = b \frac{\theta}{1-\theta} \left\{ \frac{2-2\theta}{1+\beta-2\theta} \right\}^Z \quad (Q-C) \quad (w < 0)$$

where b is defined in EQUATION 5-2, Z is the number of nearest neighbor sites, and where

$$(6-3) \quad \beta = [1 - 4\theta(1-\theta)(1 - e^{-w/kT})]^{1/2}$$

When $w = 0$ (no lateral interactions), EQUATIONS 6-1 or 6-2 reduce to the Langmuir relation, as required. If $w < 0$ (attractive interaction energies) two-dimensional condensation may set in. The occurrence of such phenomena can be detected by the presence of loops and of negative slopes over portions of the p versus θ diagram. It can be shown that condensation will occur below the critical temperature $T_c = -w/2k$ (BW) or $T_c = -w/kZ \ln \frac{Z}{Z-2} \approx -w/2.43k$ (QC) and at a pressure

$$(6-4) \quad p_c = p(1/2) = be^{zw/2kT} = \alpha e^{-(\epsilon - zw/2)/kT} \quad (w < 0)$$

for both types of approximation. The corresponding surface coverages at which condensation sets in and stops can be computed by substituting EQUATION 6-4 in 6-1 or 6-2 and solving for $\theta = \theta_c$.

At sufficiently low temperatures, condensation will occur at very low surface coverages. For such cases, $\theta_c^2 \ll 1$ and can be neglected. By expanding and solving EQUATION 6-2 with respect to θ_c one finds the approximate relation

$$(6-5) \quad \theta_c \approx \frac{-1 \pm \sqrt{1 + 4(1 + Z\eta)e^{zw/2kT}}}{2(1 + Z\eta)} \quad (w < 0)$$

where

$$(6-6) \quad \eta = 1 - e^{-w/kT} \quad (w < 0)$$

If one takes $Z = 6$, $w < 0$, and $|w| \gg kT$ the second term under the square root of (6-5)

* The present notation differs from that of Fowler and Guggenheim in that these authors denote the lateral interaction energy per pair by the product $2w/Z$, while in the present article this quantity is denoted by w . For this same reason the results quoted here often appear to be different from those given by Halsey⁶⁸ who adopted the Fowler and Guggenheim convention. The convention followed here was used by Hill,⁶⁸ except that w in Hill's paper is replaced by $-w$ here, with the understanding that $w < 0$.

In place of relation (6-2) Hill adopts the equivalent expression (see Fowler and Guggenheim⁴⁹)

$$p(\theta) = b \frac{\theta}{1-\theta} e^{zw/2kT} \left\{ \frac{(\beta - 1 + 2\theta)(1-\theta)}{(\beta + 1 + 2\theta)\theta} \right\}^{Z/2}$$

so that Hill's results differ superficially from those cited for s_1 in EQUATION 6-11.

is small. Neglect of second order quantities then yields

$$(6-7) \quad \theta_e \approx e^{Zw/2kT}$$

a relation discussed by Halsey.⁵⁵

VI A (2). *Adsorption on heterogeneous surfaces.* The adaptation of these equations to the case of heterogeneous surfaces requires a knowledge of the topographic distribution of sites associated with different energies. Such data are not available, and various simplifications must therefore be introduced. Simplifications have been suggested by Hill⁶⁸ by Halsey^{55, 57, 58, 59} and by Tompkins.¹³⁹

Hill⁶⁸ used the assumption that sites of a given energy are scattered completely at random over the surface. The lateral interactions may then be treated as though they occurred on a uniform surface. If the BW approximation is adopted, all sites associated with a given adsorption energy are covered in accordance with the relation

$$(6-8) \quad \theta_e(p, T) = \frac{1}{1 + be^{-Zw\theta/kT}/p} = \frac{1}{1 + bs_1/p}$$

Note that by contrast to the left-hand side, θ on the right-hand side refers to the coverage of the entire surface. EQUATION 6-8 resembles EQUATION 5-1 except for the presence of an exponential factor s_1 . The remarks of Section V therefore apply also to EQUATION 6-8 and, in fact, at a given T

$$(6-9) \quad \theta(p, T) = \int_0^\infty \frac{f(\epsilon) d\epsilon}{1 + \frac{\alpha}{p} e^{-Zw\theta/kT} e^{-\epsilon/kT}} = \int_0^\infty \frac{f(\epsilon) d\epsilon}{1 + \frac{\alpha}{p} s_1 e^{-\epsilon/kT}}$$

One distribution function compatible with the assumption of a random heterogeneous surface is the relation

$$(6-10) \quad \begin{cases} f(\epsilon) = 1/2\epsilon_0 & \epsilon_m - \epsilon_0 \leq \epsilon \leq \epsilon_m + \epsilon_0; \\ = 0 & \text{otherwise} \end{cases} \quad \epsilon_0 \geq 0$$

If α is assumed independent of ϵ , then use of EQUATION 6-10 in EQUATION 6-9 leads to

$$(6-11) \quad p = \frac{\alpha_0 \cdot s_1 \cdot (1 - e^{2\theta\epsilon_0/kT})(e^{-(\epsilon_m + \epsilon_0)/kT})}{(e^{(2\epsilon_0/kT)(\theta-1)} - 1)} \quad s_1 = e^{-Zw\theta/kT}$$

Note again that EQUATION 6-9 could be integrated only because the total surface coverage, θ , under the integral sign, is not a function of ϵ .

It is an easy matter to check that replacement of the BW by the QC approximation leads again to EQUATION 6-9, except that the exponential term s_1 under the integral is replaced by the quantity $s_2 = [(2 - 2\theta)/(1 + \beta - 2\theta)]^2$. Correspondingly, the term $\exp(-Zw\theta/kT)$ in EQUATION 6-11 is replaced by the new relation s_2 .

Having derived the isotherm equations for adsorption on a randomly heterogeneous surface under assumptions (a), (b), (d), (e), and (f), it is now possible

to predict the onset of two-dimensional condensation. Hill⁶⁸ has examined this point carefully by checking for the presence of negative slopes and of loops in the p versus θ diagrams plotted according to EQUATION 6-11. For details concerning the nature and types of phase changes to be expected with the two approximation methods, the reader is referred to the original article.

A second approach based on the Fowler and Guggenheim theory is attributable to Halsey.⁵⁵ This author investigated a model in which all sites associated with a given range of energies ϵ to $\epsilon + d\epsilon$ are grouped in patches so large that edge effects are avoided.

On any single patch, the condensation pressure p_i is given by EQUATION 6-4 with $\epsilon = \epsilon_i$, the adsorption energy for sites on the i th patch. This equation may be revamped as follows:

(a) Select a surface patch associated with an adsorption energy $\epsilon_i = \epsilon_L$, so that the total energy of sorption in EQUATION 6-4, inclusive of the lateral interactions, is $E_i = (\epsilon_L - Zw/2) = E_L$. Here E_L is the energy of liquefaction of the vapor. For such a patch, two dimensional condensation will occur at the saturation vapor pressure of the liquid, p_0 . Thus EQUATION 6-4 now reads $p_0 = \alpha e^{-E_L/kT}$.

(b) Consider next adsorption on patches for which $\epsilon_i = \epsilon_L + \Delta\epsilon_i$, $\Delta\epsilon_i > 0$. In that case EQUATION 6-4 reads $p_i = \alpha e^{-E_L/kT} e^{-\Delta\epsilon_i/kT}$. Division of p_i by p_0 yields

$$(6-12) \quad \frac{p_i}{p_0} = e^{-\Delta\epsilon_i/kT}; \quad \Delta\epsilon_i = -kT \ln \frac{p_i}{p_0}$$

The author next introduces the simplifying assumption (g) that for $p < p_i$, $\theta_i = 0$ and that for $p \geq p_i$, $\theta_i = 1$, i.e., that the i th patch remains empty until the critical condensation pressure (6-12) is reached, whereupon the patch becomes completely filled. The total surface coverage is then merely the sum of all patches on which condensation has taken place. These patches must have associated energies ϵ_i in the range for which $\Delta\epsilon_i \geq -kT \ln p/p_0$, p being the applied gas pressure. Accordingly,

$$(6-13) \quad \theta = \int_{-kT \ln p/p_0}^{\infty} dF(\Delta\epsilon) = \int_{-kT \ln p/p_0}^{\infty} f(\Delta\epsilon) d(\Delta\epsilon)$$

Owing to the sudden filling of each patch at a definite pressure $p = p_i$, the theory predicts the occurrence of stepwise adsorption isotherms if the distribution of energies $\epsilon_L + \Delta\epsilon$ over the patches is discontinuous.

Actually, a smoothing out of these steps will occur in most cases, but where stepwise adsorption is observed experimentally (see Section IV B) the above theory may offer an explanation in terms of this specialized model, provided $w < 0$ and $T < T_c$. The assumption $\theta_i = 0$ for $p < p_i$ and $\theta_i = 1$ for $p \geq p_i$ is only valid at temperatures well below the critical two dimensional condensation temperature T_c . Fowler and Guggenheim have demonstrated that near $T = T_c$, appreciable filling up of the patches occurs before condensation sets in, and that condensation stops long before the patch is completely covered.

If the lowest energy of condensation (exclusive of the lateral interactions) is not ϵ_L but $\epsilon_L + \Delta\epsilon_0$, $\Delta\epsilon_0 > 0$, then EQUATION 6-13 reads

$$(6-14) \quad \theta = \int_{-kT \ln p/p_0 > \Delta\epsilon_0}^{\infty} f(\Delta\epsilon) d(\Delta\epsilon) / \int_{\Delta\epsilon_0}^{\infty} f(\Delta\epsilon) d\Delta\epsilon$$

where the denominator insures normalization of f over the existing range of $\Delta\epsilon$ values, so that $\theta(p_0) = 1$.

The above theory may be further refined by taking account of multilayer condensation. One need only assume that the origin of the van der Waals forces lies one molecular diameter below the centers of the first layer adsorbate molecules. Now, from Section II B(1), one finds that ϵ , and hence $\Delta\epsilon$, fall off approximately as z^{-3} . The distribution density function for the n th layer is thus the same as for the first, except that all energies are divided by a scaling factor n^3 . Accordingly, condensation on the top of the n th layer situated at the i th patch occurs at a pressure $p_{in}/p_0 = \exp(-\Delta\epsilon_i/n^3 kT)$. In place of EQUATION 6-14 one finds^{57, 59, 60}

$$(6-15) \quad \theta_{(n)} = \int_{-n^3 kT \ln p/p_0 > \Delta\epsilon_0}^{\infty} f(\Delta\epsilon) d(\Delta\epsilon) / \int_{\Delta\epsilon_0}^{\infty} f(\Delta\epsilon) d(\Delta\epsilon)$$

To obtain an isotherm equation for multilayer formation, it is necessary to assume a dependence of f on $\Delta\epsilon$. For this purpose, Halsey introduced the trial function:

$$(6-16) \quad f = \frac{1}{\Delta\epsilon_m} e^{-\Delta\epsilon/\Delta\epsilon_m}$$

When used in EQUATION 6-15, one finds

$$(6-17) \quad \theta_{(n)} = \gamma^{n^3} e^{\Delta\epsilon_0/\Delta\epsilon_m}$$

where

$$(6-18) \quad \gamma = (p/p_0)^{kT/\Delta\epsilon_m}$$

The total surface coverage is, accordingly,

$$(6-19) \quad \theta = e^{\Delta\epsilon_0/\Delta\epsilon_m} \sum_{n=1}^{\infty} \gamma^{n^3}$$

subject to the limitation

$$(6-20) \quad \gamma^{n^3} \leq e^{-\Delta\epsilon_0/\Delta\epsilon_m}$$

in order that $\theta_{(n)} \leq 1$

Using EQUATION 6-19, Halsey derived a whole family of isotherm equations whose shape depends on the $\Delta\epsilon_0$, $\Delta\epsilon_m$, and T values selected as parameters. For details the reader is referred to his publication.⁵⁷

For certain values of the parameter in EQUATION 6-19 stepwise increments in $\theta(p, T)$ are calculated. It appears on comparison with experimental data that a large variety of the latter can be reproduced by use of EQUATION 6-19.

This in itself is of course no proof concerning the validity of the theory, for the reasons cited in Section V B. In particular, one must remember that the isotherm (6-19) is based on a very specialized patch model, on the arbitrary trial distribution (6-16), and on the use of assumptions (a), (b), (d), (e)-(g). However, together with the work of Hill, these two theories certainly represent a marked improvement over the theory of Section V A.

In this connection, there arises the question of compatibility of the assumed distribution density (6-16) with the patchwork surface model. As we have seen in Section IV, many sites are associated with a particular adsorption energy by virtue of the topographical features among which the site is located. As soon as such sites are grouped in a patch, the topographical surroundings responsible for these energy values would be missing. There is no argument concerning the fact that heterogeneous surfaces can be obtained from such patches. Polycrystalline surfaces whose sections are formed by smooth crystal planes with different Miller indices represent an example of this sort. On the other hand, it is questionable whether a *continuous* spectrum of adsorption energies is compatible with the patch model. This difficulty was evidently recognized by Halsey,⁵⁹ since he states that EQUATION 6-16 represents an arbitrary distribution density and "that the type of heterogeneity postulated is not that of 'active points' in the surface layers. Rather than that, it is suggested that different faces of a polycrystalline nature combine in various ways to produce heterogeneity." The manner in which this process is presumed to occur is not clarified further.

Halsey's theory has been criticized by Hall,⁵³ who states that surface sites "have a considerably greater attractive energy than the underlying layers, and if this is the case, the attractive energy should be closely approximated by integration over the surface layer only." Such a hypothesis would demand that ϵ decrease as z^{-4} . This argument may be questioned on several points.^{59, 60} In the first place, heterogeneity in physical adsorption results from differences in the arrangement of surface atoms relative to the interior of the lattice, or from surface topography. There is thus no reason to endow surface atoms with special properties in comparison to which the bulk effects of the crystal can be neglected. Furthermore, at large distances from the surface the London relation (2-17) should become applicable, which requires that ϵ be proportional to z^{-3} .

Halsey⁵⁹ did amend his original theory, however, so as to include a term depending on z^{-4} as well as a term to take repulsion potentials into account. These refinements will not be discussed further.

Still another approach was used by Tompkins.¹³⁹ He suggested that, at high surface coverage, each molecule is confined in its movement to a two-dimensional cell of area **A**. One can show by statistical methods that if sites of equal adsorption energy ϵ_i are grouped in patches then the isotherm for the *i*th patch reads

$$(6-21) \quad p = K b e^{-\epsilon_i/kT} \frac{\theta_i}{(1 - \theta_i^{1/2})^2} e^{\theta_i^{1/2}/(1 - \theta_i^{1/2})}$$

where K is a constant. This equation must be solved for $\theta_\epsilon(p, T)$; it can then be integrated to yield $\theta(p, T) = \int_0^\infty \theta_\epsilon dF(\epsilon)$. Various approximation methods for solving EQUATION 6-21 for θ_ϵ are discussed. All of these lead to Langmuir type relations.

VI B. Variable lateral interactions. The treatment of Section VI A is based on the fairly artificial assumptions (e) and (f). In reality, the interaction energy varies continually with changes in distance between adsorbed atoms.

One method of treating this type of interaction is to regard the adsorbate as an imperfect two-dimensional gas. Such a treatment also eliminates the necessity of assumption *d* of Section V B (1) pertaining to localized adsorption. This theory should be applicable to isotherms obtained at higher temperatures.

VI B (1). Adsorption on uniform surfaces. Numerous papers have dealt with the theory of adsorption of two-dimensional gases on uniform surfaces. The contributions up to 1942 are reviewed by Brunauer.¹⁹ Other work since that time includes contributions by Everett,⁴⁶ Vol'kenshtein,¹⁴³ Hill,^{65, 68} de Boer,³⁶ and Barrer and Robins.⁶ Correlation procedures have been developed^{46, 143, 155} which enable one to determine the equation of state or lateral interaction law for the adsorbate from adsorption isotherms or vice versa.

In recent years the two dimensional van der Waals equation

$$(6-22) \quad (\phi + a_2/A^2)(A - b_2) = RT$$

has been used extensively as an equation of state. Here a_2 and b_2 are the two-dimensional analogues of the three-dimensional van der Waals constants; ϕ is the spreading pressure and A the surface area. The corresponding adsorption isotherm can be shown to be^{36, 66}

$$(6-23) \quad \begin{aligned} p(\theta) &= b \frac{\theta}{1 - \theta} e^{\theta/(1-\theta) - 2a_2\theta/\sigma kT} \\ &= k_2 \frac{\theta}{1 - \theta} e^{[\theta/(1-\theta) - k_1(a_2, b_2) \cdot \theta]} \end{aligned}$$

where σ is the effective cross section of the adsorbate on the surface.

In his monograph, de Boer³⁶ discusses the use and properties of EQUATION 6-23 extensively. The author takes the view that the establishment of equilibrium conditions between the gas phase and the adsorbed phase always necessitates surface mobility. He therefore prefers the use of EQUATION 6-23 to use of the Langmuir relation (5-1) or of the modified Langmuir relations (6-1 and 6-2) pertaining to localized adsorption. He shows that the various isotherms observed experimentally can be reproduced by EQUATION 6-23 with proper selection of the constants k_1 and k_2 , k_1 depending only on the properties of the adsorbate. As k_1 passes a certain critical value, loops are obtained in the θ versus p plot indicative of the occurrence of two-dimensional condensation. (See also Hill⁶⁸.) For relatively small values of k_2 (~ 0.1) and relatively large values of k_1 (~ 10), representing conditions of strong lateral interactions but weak adsorbent-adsorbate binding, condensation may set in at very low pressures

and surface coverage, and may end at θ values in the range $0.4 < \theta < 0.8$ depending on the numerical values selected for k_1 and k_2 . It is repeatedly emphasized in this treatment that differentiation between the Langmuir isotherms and EQUATION 6-23 may be a very difficult matter.

Theimer¹³⁵ has suggested the use of another equation of state which leads to an adsorption isotherm resembling EQUATION 6-23. Barrer and Robins⁶ have extended EQUATION 6-23 to include multilayer sorption.

VI B (2). Adsorption on heterogeneous surfaces. If surfaces are not homogeneous, one can apply EQUATION 6-23 separately to all portions of a surface associated with a range of adsorption energies ϵ to $\epsilon + d\epsilon$ or k_2 to $k_2 + dk_2$. For every surface segment, the equilibrium gas pressure is substituted on the left, and θ_ϵ (θ in EQUATION 6-23) is obtained on the right by use of an approximation method. The total surface coverage is then found from

$$(6-24) \quad \theta(p, T) = \int_0^\infty \theta_\epsilon dF(\epsilon)$$

as usual. Since the adsorbate is mobile, $F(\epsilon)$ or $F(k_2)$ represents a distribution function over all accessible parts of the surface; by contrast, $F(\epsilon)$, in Section V (localized adsorption) referred to a distribution in ϵ over "adsorption sites", *i.e.* over locations of potential minima on the surface. Use of an assumed relation $dF(\epsilon)$ in the integral then yields θ .

The results of some numerical computations by de Boer³⁶ have been cited in his monograph. Depending on the selection of the parameters k_1 and k_2 and on the choice of a trial function $F(\epsilon)$, many different types of isotherms can be generated. If $F(\epsilon)$ is sufficiently discontinuous and $k_1 < k_{1c}$, $T < T_c$ stepwise isotherms are obtained, indicative of two dimensional condensation over portions of the surface. If a surface is characterized by a continuous variation in k_2 or ϵ , a gradual change in slope of the θ versus p curves is found, and all of the five types of isotherms of the Brunauer classification¹⁹ can be reproduced. The application of these methods to a description of multilayer formation is also given in the monograph.

Tompkins¹³⁹ has considered the adsorption of gases under a variety of conditions. For example, if nonlocalized adsorption occurs and the adsorbate is in a highly compressed state, the isotherm equation for a uniform surface reads (approximately)

$$(6-25) \quad p = \frac{b_\epsilon}{\sigma} \frac{\theta_\epsilon}{1 - 2\theta_\epsilon} e^{2\theta_\epsilon/(1-2\theta_\epsilon)}$$

If the surface consists of patches of equal energy, one can solve for θ_ϵ on the right by approximation methods. One can then find θ by use of EQUATION 6-24. Tompkins derived an isotherm equation in the limit of low coverage and under the assumption of an exponential distribution of energies over the patches. He obtained a Freundlich isotherm under such conditions.

Again, assuming a patch model surface, Tompkins derived EQUATION 6-23 for a two-dimensional van der Waals gas. He showed that if θ_ϵ is small, EQUATION

TION 6-23 can be rewritten in the form

$$(6-26) \quad \theta_{\epsilon} \approx \frac{1}{3 + \frac{1.3b}{p}}$$

This is in a form similar to EQUATION 5-1. Thus, if EQUATION 6-26 is used in 6-24 under the assumption that f depends exponentially on ϵ the Freundlich isotherm equation is again found.

For the case where adsorption sites are randomly distributed over the surface, Tompkins derived an equation of the form

$$(6-27) \quad p = \frac{b}{\sigma} \frac{\theta_{\epsilon}}{1 - 2\theta} e^{2\theta_{\epsilon}/(1-2\theta)}$$

Note that both the coverage of the entire surface and the occupancy of portions associated with energy ϵ_i is involved here. Again, one must first solve for θ_{ϵ} and then use EQUATION 6-24. The author discusses analytical solutions for several limiting cases.

The feature common to all the above calculations is the setting up of a functional relation $p(\theta_{\epsilon})$. This relation is then solved by approximation methods or numerically for $\theta_{\epsilon}(p)$. The result is then used in EQUATION 6-24, in conjunction with a trial function F , to obtain θ .

Tompkins showed by use of approximation procedures that the isotherms obtained here are functionally very closely related to those treated by the methods of Section V A. This is another instance of the difficulty so frequently encountered in the theory of gas adsorption: namely, that, under normal conditions, the adsorption data alone do not suffice to distinguish between several possible adsorption mechanisms.

VI C. Heterogeneity versus lateral interactions. It has been shown that a given isotherm equation could be correlated either with a distribution of adsorption energies under assumption (c) (Section IV A and B) or with an equation of state for a nonideal adsorbate on a smooth surface [Section VI B (1)]. There has been much discussion in the literature as to which view is preferable.

Isolated instances undoubtedly exist in which the adsorption is predominantly of one type. The work of Rhodin^{104, 105} described in Section IV B represents a case where the smooth surface - lateral interactions model seems applicable. In general, however, neither surface inhomogeneity nor lateral interactions can be ignored.

As Beeck¹² states: "Any attempt to establish whether a surface is homogeneous or heterogeneous must necessarily be futile because no truly homogeneous surface exists." Furthermore, the types of forces responsible for the attraction of gases to surfaces of solids are also operative among the adsorbed molecules themselves. The theories of the type discussed in Sections VI A and B therefore seem to be of the greatest utility.

VII. Surface Heterogeneity in Chemisorption

We conclude with a few remarks concerning the concept of surface heterogeneity in chemisorption and catalysis.

Taylor^{125, 126, 127, 130} introduced the theory of "active centers" to explain why only small portions of the surface of a solid are active in catalysis or why the presence of small quantities of poison readily inhibit catalytic processes. He postulated that, at different geometric locations, the number of "free valencies" varies, thus giving rise to gradations in the energies of chemisorption. The number of surface atoms active in catalysis determines the sensitivity of the solid to poisons. This concept has been used repeatedly to explain the drop in the heats of chemisorption with surface coverage.

A related problem concerns the variation in the rates of chemisorption with surface coverage. Taylor postulated the existence of a distribution in activation energies^{128, 129} of adsorption over the surface to explain these variations.

The interpretation of experimental data in terms of the above concepts has been discussed in numerous publications;^{23, 38, 54, 56, 58, 82, 109, 110, 112, 118, 134} in fact, much of the theory of Section V A was developed for application to chemisorption phenomena.

There has been a tendency lately to reinterpret the above concepts. Before going into this problem, it should be noticed that the types of surface heterogeneities present in physical sorption cannot be used in explaining chemisorption data; for the enthalpy changes in activated sorption are far larger than those encountered in physical adsorption, so that the relatively minor variations in physical sorption potentials can have only a secondary effect on the overall binding energy. Likewise, lateral repulsions between the chemisorbed species are often insufficient¹⁰⁷ to explain the observed drop of ΔH_s with increasing θ . This raises the question whether the theories discussed in monographs by Roberts¹⁰⁸ and Miller⁹⁵ are really generally applicable.

The existence of active centers can be explained by the presence of impurities on the surface.¹³¹ It is well known that such surface phenomena as work functions of metals are greatly modified by the presence of impurities. Taylor suggests that, owing to the valency induction effect,¹¹⁶ the impurities deposited on a surface tend to assume "anomalous valencies" and that this is reflected in the occurrence of abnormal activation energies of chemisorption.

Cases where the equilibration between adsorbate and adsorbent requires considerable time are frequently explained in terms of other phenomena accompanying chemisorption, such as solution of the adsorbate in the metal lattice^{9, 10, 11, 12, 122, 146} or diffusion processes.¹⁴⁸

Some of the more recent theories explicitly consider the alteration of the surface structure of the adsorbent in the chemisorption process. Vol'kenshtein¹⁴⁴ considered lattice defects on the surface as adsorption centers in chemisorption. These defects may occur at 0° K. in the lattice (biographical defects) and/or they may be created thermally at higher temperatures. If only one type of defect exists, the surface is energetically homogeneous, but with two or more types of defects, heterogeneity effects can occur. In Vol'kenshtein's theory, these centers can move about, and can interact so as to annihilate each other or to produce new defects. The number of adsorption sites need not necessarily remain constant. New centers of adsorption may be created in the chemisorption process itself. Using such a model, Vol'kenshtein was able to

predict a decrease in ΔH_s with increasing surface coverage and to explain changes in the rates of adsorption.*

As an example of what at first might seem to be an abstract model, one might consider the ionic lattice $M^+ A^-$ containing an impurity I located over some of the anions A^- . As T is raised, the impurity I may become ionized and the liberated electron may then migrate to an adjacent site, converting M^+ into M . Suppose the M atom represents a chemisorption site; then it is clear that, as a result of thermal disorders, adsorption centers are "created" according to the "net reaction" $M^+ + I = M + I^+$. Furthermore, unless this reaction is practically instantaneous with respect to the rate of chemisorption, the number of adsorption centers changes with θ . Also, the rate of chemisorption will depend on the rate of production of the centers if the latter step is slow. Finally, the heats of adsorption can be calculated by noting that the creation of every adsorption site requires an energy expenditure, E , and that the energy $-\epsilon$ is liberated upon occupancy of every such site. Owing to the fact that the number of sites available for chemisorption decreases with θ , the variation of ΔH_s with θ is explained by this model. Other detailed discussions are presented, by Cabrera,²⁰ Taylor,¹³² and in an extension of the original treatment, by Boudart and Taylor.¹⁵

A somewhat similar approach, utilizing semiconductor theory has been developed recently.^{1, 16, 39, 61, 147} Semiconductors may contain quasi-free electrons or holes, and when a gas comes in contact with the adsorbent the possibility for a charge transfer between the two species exists. This results in the formation of bonds ranging through intermediates from covalent to ionic types.

As a result of the transfer, a space charge is generated in the solid which may reach many atomic layers into the bulk phase of the semiconductor. This space charge acts as a potential barrier, impeding and ultimately preventing further electron transfers. At that point, chemisorption stops. The higher the barrier, the less is the heat of chemisorption, which ultimately drops to zero. By considering the diffusion rates of charge carriers to the surface one can calculate the rates of chemisorption.

It is seen that all these phenomena can be accounted for in terms of a theory detailing the electronic properties of semiconductors relative to those of the adsorbate species. In the chemisorption process itself the semiconductor properties of the solid are changed and this accounts for effects previously described in terms of pre-existing energy heterogeneities. This theory does not rule out the presence of active centers on the surface. It does show that the observed chemisorption phenomena can be explained in terms of a quasi-heterogeneity, induced on the surface owing to the adsorption process itself. In this view, then, an adsorbate does more than just cover a surface site; it alters the nature of the surface layers and thus influences the entire remaining course of chemisorption. These concepts are explained in full detail by Boudart.¹⁶

Boudart¹⁷ has used a similar model to explain the apparent heterogeneity observed in chemisorption of gases on metals. As soon as a metallic surface

* Translations of this article are available upon payment of a small fee from the National Research Council of Canada, National Research Laboratories, Division of Information Services, Ottawa, Canada.

is covered appreciably, an electric double layer is formed, owing to the electron transfer during the establishment of a metallic surface bond. This barrier changes the work function of the metal and, correspondingly, the heat of chemisorption. Once again, the fall in ΔH_s with θ is explained in terms of physical changes in the adsorbent-adsorbate system. Naturally, repulsive lateral interactions or pre-existing heterogeneities (owing to the presence of impurities on the surface, for example) may also contribute to the variation of ΔH_s with θ . However, they are not essential in the explanation of these effects. Beeck¹² postulates a similar mechanism, in terms of a lowering of surface energy of sites adjacent to occupied sites.

Finally, some theoretical studies by H. A. Taylor and Thon¹²⁴ should be mentioned, concerning the rates of chemisorption. Very often these rates obey the empirical law $\frac{du}{dt} = ae^{-\alpha n}$, α a constant. These authors cite evidence that such a rate law can be interpreted as an explosive production of sites which takes place when the gas comes in contact with the adsorbent. This is then followed by a slower, spontaneous disappearance of sites, resulting in a drop in the rate of chemisorption. Experimental evidence in favor of the annihilation of such sites is reviewed.

Acknowledgments

The writer wishes to express his gratitude to Doctor M. Boudart for many fruitful discussions.

VIII. Bibliography

1. AIGRAIN, P. & C. DUGAS. 1952. *Z. Elektrochem.* **56**: 363.
2. ASHWORTH, F. 1951. *Advances in Electronics*. Academic Press Inc. New York, N. Y. **3**: 1.
3. BARDEEN, J. 1940. *Phys. Rev.* **58**: 727.
4. BARRER, R. M. 1937. *Proc. Roy. Soc. (London, England)* **A161**: 476.
5. BARRER, R. M. 1950. *Disc. Faraday Soc.* **8**: 88.
6. BARRER, R. M. & A. B. ROBINS. 1951. *Trans. Faraday Soc.* **47**: 773.
7. BECKER, J. A. 1951. *Bell System Tech. J.* **30**: 907.
8. BEEBE, R. A. 1943. *Handbuch der Katalyse*. Springer. Vienna, Austria **4**: 473.
9. BEECK, O., A. E. SMITH & A. WHEELER. 1940. *Proc. Roy. Soc. (London, England)* **A 177**: 62.
10. BEECK, O., A. W. RITCHIE & A. WHEELER. 1948. *J. Colloid Sci.* **3**: 505.
11. BEECK, O., J. W. GIVENS & A. W. RITCHIE. 1950. *J. Colloid Sci.* **5**: 141.
12. BEECK, O. 1950. *Advances in Catalysis*. Academic Press, Inc. New York, N. Y. **2**: 151.
13. BORN, M. & J. E. MAYER. 1932. *Z. Phys.* **75**: 1.
14. BORN, M. 1921. *Atomtheorie des festen Zustandes*. 2nd ed. Leipzig, Germany.
15. BOUDART, M. & H. S. TAYLOR. 1952. *I. Farkas Memorial Volume, Research Council of Israel, Jerusalem*. Special Publication No. **1**: 223. Oth Printing Press. Haifa, Israel.
16. BOUDART, M. 1952. *J. Am. Chem. Soc.* **74**: 1531.
17. BOUDART, M. 1952. *J. Am. Chem. Soc.* **74**: 3556.
18. BRUNAUER, S., P. H. EMMETT & E. TELLER. 1938. *J. Am. Chem. Soc.* **60**: 309.
19. BRUNAUER, S. 1943. *The Adsorption of Gases and Vapors*. Princeton University Press. Princeton, N. J.
20. CABRERA, N. 1950. *Z. Elektrochem.* **56**: 294.
21. CABRERA, N. & N. F. MOTT. 1949. *Rep. Prog. Phys.* **12**: 163.
22. CASSIE, A. B. D. 1945. *Trans. Faraday Soc.* **41**: 450.
23. CONSTABLE, F. H. 1925, 1926. *Proc. Cambridge Phil. Soc.* **22**: 738; **23**: 172, 593.

24. COOK, M. A., D. H. PACK & A. G. OBLAD. 1951. *J. Chem. Phys.* **19**: 367.
25. CRAWFORD, V. A. & F. C. TOMPKINS. 1948. *Trans. Faraday Soc.* **44**: 698.
26. CREMER, E. & G. M. SCHWAB. 1929. *Z. physik. Chem. A.* **144**: 243; *B.* **5**: 406.
27. CREMER, E. & S. FLÜGGE. 1938. *Z. physik. Chem. B.* **41**: 453.
28. CREMER, E. 1947. *Monatsh. f. Chem.* **77**: 126.
29. CREMER, E. 1949. *J. chim. phys.* **46**: 411.
30. CROWELL, A. D. 1952. *Am. J. Phys.* **20**: 89.
31. DAMKÖHLER, G. 1933. *Z. physik. Chem. B.* **23**: 58.
32. DE BOER, J. H. & J. F. H. CUSTERS. 1934. *Z. physik. Chem. B.* **25**: 225.
33. DE BOER, J. H. 1936. *Trans. Faraday Soc.* **32**: 10.
34. DE BOER, J. H. 1935. *Electron Emission and Adsorption Phenomena*. Cambridge University Press. London, England.
35. DE BOER, J. H. 1950. *Advances in Colloid Science*. Interscience Publishers. New York. **3**: 1 ff.
36. DE BOER, J. H. 1953. *The Dynamical Character of Adsorption*. The Clarendon Press. Oxford, England.
37. DE BOER, J. H., & S. KRUYER. 1952, 1953. *Proc. Konink. Nederland. Akad. v. Wetenschap. B* **55**: 451; *B* **56**: 67.
38. DE BRUIJN, H. 1950. *Disc. Faraday Soc.* **8**: 69.
39. DOWDEN, D. A. 1950. *J. Chem. Soc.* : 242.
40. DRAIN, L. E. & J. A. MORRISON. 1952. *Trans. Faraday Soc.* **48**: 316.
41. DRAIN, L. E. & J. A. MORRISON. 1953. *Trans. Faraday Soc.* **49**: 654.
42. DRAIN, L. E. 1953. *Trans. Faraday Soc.* **49**: 650.
43. DRENAN, J. W. & T. L. HILL. 1949. *J. Chem. Phys.* **17**: 775.
44. DUBININ, M. & M. CHRENOVA. 1936. *Zhur. Fiz. Khim.* **9**: 1204.
45. DUNKEL, M. & H. MARK. 1929. *Monatsh.* **53/54**: 764.
46. EVERETT, D. H. 1950. *Trans. Faraday Soc.* **46**: 942.
47. FORESTI, B. 1929. *Gazz. Chim. Ital.* **59**: 243.
48. FOWLER, R. H. 1936. *Statistical Mechanics* : 832. 2nd ed. MacMillan. New York, N. Y.
49. FOWLER, R. H. & E. A. GUGGENHEIM. 1939. *Statistical Thermodynamics*. Cambridge University Press. Cambridge, England.
50. FRICKE, R. 1943. *Handbuch der Katalyse*. Springer. Vienna, Austria **4**: 1.
51. GLASSTONE, S., K. LAIDLER & H. EYRING. 1941. *Theory of Rate Processes*. McGraw-Hill, New York, N. Y.
52. GOMER, R. & D. A. SPEER. 1953. *J. Chem. Phys.* **21**: 73.
53. HALL, W. K. 1952. *J. Am. Chem. Soc.* **74**: 1081.
54. HALSEY, G. D., JR. & H. S. TAYLOR. 1947. *J. Chem. Phys.* **15**: 624.
55. HALSEY, G. D., JR. 1948. *J. Chem. Phys.* **16**: 931.
56. HALSEY, G. D., JR. 1949. *J. Chem. Phys.* **17**: 758.
57. HALSEY, G. D., JR. 1951. *J. Am. Chem. Soc.* **73**: 2693.
58. HALSEY, G. D., JR. 1951. *J. Phys. & Colloid Chem.* **55**: 21.
59. HALSEY, G. D., JR. 1952. *J. Am. Chem. Soc.* **74**: 1082.
60. HALSEY, G. D., JR. 1952. *Advances in Catalysis*. Academic Press, Inc. New York, N.Y. **4**: 259.
61. HAUFFE, K. & H. J. ENGELL. 1952. *Z. Elektrochem.* **56**: 366.
62. HELLMANN, H. 1937. *Einfuehrung in die Quantenchemie*. Deuticke. Leipzig, Gernay, & Vienna, Austria.
63. HERZFELD, K. K. 1929. *J. Am. Chem. Soc.* **51**: 2608.
64. HILL, T. L. 1946. *J. Chem. Phys.* **14**: 263.
65. HILL, T. L. 1946. *J. Chem. Phys.* **14**: 441.
66. HILL, T. L. 1947. *J. Chem. Phys.* **15**: 767.
67. HILL, T. L. 1948. *J. Chem. Phys.* **16**: 181.
68. HILL, T. L. 1949. *J. Chem. Phys.* **17**: 762.
69. HILL, T. L. 1952. *Advances in Catalysis*. Academic Press, Inc. New York, N. Y. **4**: 211.
70. HONIG, J. M. & L. H. REYERSON. 1952. *J. Phys. Chem.* **56**: 140.
71. HONIG, J. M. 1953. *J. Phys. Chem.* **57**: 349.
72. HÜCKEL, E. 1928. *Adsorption und Kapillarkondensation*. Akademische Verlagsgesellschaft. Leipzig, Germany. 126 ff.
73. ILIN, B. 1924, 1925. *Phil. Mag.* **48**: 193; **50**: 1114.
74. ILIN, B. 1925. *Z. Physik.* **33**: 435.
75. JACQUET, E. 1925. *Fortschr. Chem. Physik u. physik. Chem. B.* **18**: H 7.
76. JURA, G. & D. CRIDDLE. 1951. *J. Phys. & Colloid Chem.* **55**: 163.

77. KAR, S. C. 1925. *Physik. Z.* **26**: 615.
78. KAR, K. & A. GANGULI. 1928. *Physik. Z.* **30**: 918.
79. KEMBALL, C. 1950. *Advances in Catalysis*. Academic Press, Inc. New York, N. Y. **2**: 233.
80. KIRKWOOD, J. G. 1932. *Physik. Z.* **33**: 57.
81. KOCHENDÖRFER, A. 1952. *Z. Elektrochem.* **56**: 283.
82. LANGMUIR, I. 1918. *J. Am. Chem. Soc.* **40**: 1361.
83. LENEL, F. V. 1933. *Z. physik. Chem. B.* **23**: 379.
84. LENNARD-JONES, J. E. 1931. *Proc. Phys. Soc.* **43**: 471.
85. LENNARD-JONES, J. E. 1932. *Trans. Faraday Soc.* **28**: 333.
86. LENNARD-JONES, J. E. & B. M. DENT. 1928. *Trans. Faraday Soc.* **24**: 92.
87. LONDON, F. 1930. *Z. Physik.* **63**: 245.
88. LONDON, F. 1930. *Z. physik. Chem. B.* **11**: 222.
89. LORENZ, R. & A. LANDÉ. 1922. *Z. anorg. Chem.* **125**: 47.
90. MADELUNG, E. 1918. *Physik. Z.* **19**: 524.
91. MAGNUS, A. 1922. *Z. physik. Chem.* **142**: 401.
92. MARGENAU, H. 1939. *Rev. Modern Phys.* **11**: 1.
93. MARGENAU, H. & W. G. POLLARD. 1941. *Phys. Rev.* **60**: 128.
94. McMILLAN, W. G. 1947. *J. Chem. Phys.* **15**: 390.
95. MILLER, A. R. 1949. *The Adsorption of Gases on Solids*. Cambridge University Press. Cambridge, England.
96. ORR, W. J. C. 1939. *Trans. Faraday Soc.* **35**: 1247.
97. PEASE, R. N. 1923. *J. Am. Chem. Soc.* **45**: 2296.
98. POHL, R. W. 1952. *Naturwissenschaften* **39**: 9.
99. POLANYI, M. & F. LONDON. 1930. *Naturwissenschaften* **18**: 1099.
100. POLANYI, M. 1932. *Trans. Faraday Soc.* **28**: 316.
101. POLLEY, M. H., W. D. SCHAEFFER & W. R. SMITH. 1953. *J. Phys. Chem.* **57**: 469.
102. PROSEN, E. J. R., R. G. SACHS & E. TELLER. 1940. *Phys. Rev.* **57**: 1066.
103. PROSEN, E. J. R. & R. G. SACHS. 1942. *Phys. Rev.* **61**: 65.
104. RHODIN, T. N. 1950. *J. Am. Chem. Soc.* **72**: 5691.
105. RHODIN, T. N. 1953. *J. Phys. Chem.* **57**: 143.
106. RIDEAL, E. K. 1930. *An Introduction to Surface Chemistry*. Cambridge University Press. London, England.
107. RIDEAL, E. K. & B. M. W. TRAPNELL. 1950. *J. chim. phys.* **47**: 126.
108. ROBERTS, J. K. 1939. *Some Problems in Adsorption*. Cambridge University Press. Cambridge, England.
109. ROGINSKII, C. Z. 1944. *Doklady Akad. Nauk. S. S. S. R.* **45**: 61, 194.
110. ROGINSKII, C. Z. 1945. *Doklady Akad. Nauk. S. S. S. R.* **47**: 412; 478; 558.
111. ROGINSKII, C. Z. & O. TODES. 1946. *Acta Physicochim. U. R. S. S.* **21**: 519.
112. ROGINSKII, C. Z. 1945, 1947. *Acta Physicochim. U. R. S. S.* **20**: 227; **22**: 61.
113. RUSHBROOKE, G. S. 1949. *Introduction to Statistical Mechanics*. Oxford University Press. Oxford, England.
114. SCHILLING, K. 1947. *Monatsh.* **77**: 134.
115. SEITZ, F. 1952. *Imperfections in Nearly Perfect Crystals*. National Research Council. Wiley. New York, N. Y. : 3
116. SELWOOD, P. W. 1949. *Bull. soc. chim. France.* : D 167.
117. SEXL, T. 1928. *Z. physik.* **48**: 607.
118. SHERMAN, A. & H. EYRING. 1932. *J. Am. Chem. Soc.* **54**: 2661.
119. SIPS, R. 1948. *J. Chem. Phys.* **16**: 490.
120. SIPS, R. 1950. *J. Chem. Phys.* **18**: 1024.
121. SLATER, J. C. & J. G. KIRKWOOD. 1931. *Phys. Rev.* **37**: 682.
122. STEACIE, E. W. 1932. *Trans. Faraday Soc.* **28**: 617.
123. STRANSKI, I. N. 1930. *Z. Elektrochem.* **36**: 25.
124. TAYLOR, H. A. & N. THON. 1952. *J. Am. Chem. Soc.* **74**: 4169.
125. TAYLOR, H. S. 1925. *Proc. Roy. Soc. (London, England) A.* **108**: 105.
126. TAYLOR, H. S. 1926. *J. Phys. Chem.* **30**: 145.
127. TAYLOR, H. S. 1929. *Z. Elektrochem.* **35**: 542.
128. TAYLOR, H. S. 1931. *J. Am. Chem. Soc.* **53**: 578.
129. TAYLOR, H. S. 1931. *Colloid Chemistry*. Chemical Catalog Company, Inc. New York, N. Y. **3**: 107 ff.
130. TAYLOR, H. S. 1934. *Acta Physicochim. U. R. S. S.* **1**: 397.
131. TAYLOR, H. S. 1950. *Disc. Faraday Soc.* **8**: 9.
132. TAYLOR, H. S. 1950. *Annual Review of Physical Chemistry*. Annual Reviews, Inc. Stanford, Calif. **1**: 271 ff.

133. TEMKIN, M. & V. LEVICH. 1946. *Zhur. Fiz. Khim.* **20**: 1441.
134. TEMKIN, M. 1948. *Uspekhi Khim.* **17**: 174.
135. THEIMER, O. 1951. *Z. Elektrochem.* **55**: 709.
136. TITCHMARSH, E. C. 1948. *The Theory of Fourier Integrals*. 2nd ed. Oxford University Press. Oxford, England.
137. TODES, O. M. & A. K. BONDAREVA. 1948. *Zhur. Priklad. Khim.* **21**: 693.
138. TOLPIN, J. G., G. S. JOHN & E. FIELD. 1953. *Advances in Catalysis*. Academic Press, New York, N.Y. **5**: 217.
139. TOMPKINS, F. C. 1950. *Trans. Faraday Soc.* **46**: 569.
140. VAN ARKEL, A. E. & J. H. DE BOER. 1936. *La Valence et l'Electrostatique*. Librairie Alcan. Paris, France. Chapter 11.
141. VERWEY, E. J. W. & J. H. DE BOER. 1940. *Rec. trav. chim.* **59**: 633.
142. VERWEY, E. J. W. 1946. *Rec. trav. chim.* **65**: 521.
143. VOL'KENShteIN, F. F. 1947. *Zhur. Fiz. Khim.* **21**: 163.
144. VOL'KENShteIN, F. F. 1949. *Zhur. Fiz. Khim.* **23**: 917.
145. WALKER, W. C. & A. C. ZETTMAYER. 1948. *J. Phys. & Colloid Chem.* **52**: 47, 58.
146. WARD, A. F. H. 1932. *Trans. Faraday Soc.* **28**: 399.
147. WEISZ, P. B. 1952, 1953. *J. Chem. Phys.* **20**: 1483; **21**: 1531.
148. WICKE, E. 1949. *Z. Elektrochem.* **53**: 279.
149. WIDDER, D. V. 1946. *The Laplace Transform*. Princeton University Press. Princeton, N. J.
150. WIG, E. O. & S. B. SMITH. 1951. *J. Phys. and Colloid Chem.* **55**: 27.
151. WILKINS, F. J. 1938. *Proc. Roy. Soc. (London, England) A.* **164**: 496, 510.
152. WOLF, P. M. & N. RIEHL. 1932. *Z. anorg. Chem.* **45**: 400.
153. YOUNG, D. M. 1951. *Trans. Faraday Soc.* **47**: 1228.
154. ZELDOWITSH, J. 1934. *Acta Physicochim. U. R. S. S.* **1**: 961.
155. ZHUKHOVITSKII, A. A. & R. N. RUBINSTEIN. 1949. *Doklady Akad. Nauk. S. S. S. R.* **65**: 41.

THE KINETICS OF SURFACE PROPERTIES

By H. Austin Taylor

New York University, New York, N. Y.

It is customary at the present time to divide any discussion of reaction kinetics into two classes: homogeneous and heterogeneous, with the inference that what may be true for the one class need not be true for the other. From the point of view that, in each class, there is a chemical reaction between the same or different species it is apparent that a single basis must underlie both classes, and that each must be susceptible to the same basic kinetic principles. There cannot exist any chasm between the two, and apparent discrepancies which seem to appear must be the sign posts to a correct unification.

Discrepancies between initial rate and course of reaction kinetic laws in heterogeneously catalyzed systems have recently been pointed out¹⁶ for several reactions. These discrepancies have been interpreted to indicate a far more significant role of the surface than is at present generally supposed. A further analysis of some of the reactions which are now "classical" and accepted as convincing evidence for the current point of view demonstrates the necessity for a searching review of the whole basis of present-day concepts.

The decomposition of nitrous oxide on a platinum surface was interpreted by Hinshelwood⁵ as a reaction of first order in nitrous oxide inhibited by the oxygen produced. The course of the reaction was expressed by the relation: $-d(\text{N}_2\text{O})/dt = k(\text{N}_2\text{O})/[1 + b(\text{O}_2)]$, and at 911° C. the values of k and b were found to be 0.00647 and 0.0341 respectively. The existence of the oxygen inhibition was further tested by measuring initial rates of decomposition of various mixtures of nitrous oxide and oxygen. The latter data, expressed as the per cent of reaction r in one minute, fitted the relation: $r = 64/[1 + 0.02(\text{O}_2)]$. The similarity of the form of this relation to the course of reaction relation above is the evidence for the validity of the suggested mechanism. The values of k and b found using nitrous oxide alone, however, would yield an initial rate expression $r = 39/[1 + 0.03(\text{O}_2)]$ which is only about one half that found for mixtures with oxygen. Where is the inhibition? Rather would it seem that oxygen has accelerated the decomposition. Steacie¹² finds for N_2O decomposition on spongy platinum that added oxygen has no effect on the decomposition but still attributes the decrease in rate to inhibition by oxygen produced in the reaction.

Schwab, Staeger, and Baumbach⁹ studied the decomposition of nitrous oxide on a number of oxides. The results are interpreted as indicating a first order reaction without inhibition on MgO , CaO , SrO , Al_2O_3 and La_2O_3 ; an oxygen inhibition on CdO , CuO and NiO and self inhibition, that is, by nitrous oxide on In_2O_3 . It is claimed that the inhibitions disappear at sufficiently high temperature and the reactions become simple first order.

Still more recently Haufler, Glang, and Engel⁴ have used the decomposition of nitrous oxide mixed with air in a flow system, to study the influence of electron defects in oxide catalysts. They have determined the order of the reaction by finding what function of the fractional decomposition in a given time

TABLE 1
APPROXIMATE ORDER OF DECOMPOSITION OF NITROUS OXIDE

Catalyst	Order
NiO	3/2
NiO + 0.1% Li ₂ O	3/4
NiO + 0.5% Li ₂ O	1
NiO + 1.0% Li ₂ O	3/2
NiO + 3.0% Li ₂ O	2
NiO + 1.0% In ₂ O ₃	1
CuO	3/2
CuO + 1% Cr ₂ O ₃	3/2
ZnO	3/2
Cr ₂ O ₃	3/2

will yield a linear plot for the Arrhenius equation. Disregarding the question of the validity of this procedure, the approximate orders found are given in TABLE 1.

If one were to assume in heterogeneous reactions, as is commonly done in homogeneous systems, that the order of a reaction can be determined either by comparing initial rates of separate runs, or by following the progress of the reaction throughout a run, or by determining the fractional decomposition in a given time, one would be presented with a conflict of orders for the nitrous oxide decomposition that would tax the imagination for mechanistic interpretations. The data show very definitely that the specific rate falls during the course of a run. This is inhibition. The only question is: what is the cause of the inhibition? The assumptions offered so far include the obvious: oxygen and nitrous oxide. There has been a complete disregard of the role of the surface. The experimental data by Hinshelwood can be accounted for with equal adequacy by an exponential rate decrease attributable to a decay of active centers on the surface.

The dehydration of isopropyl alcohol on bauxite was found by Dohse and Kälberer³ to be first order during the course of a run but zero order for the initial rates from run to run. The first order was attributed to a poisoning of the aluminum oxide by water. To test this assumption, barium oxide was added to the bauxite, in the expectation that barium oxide would remove the water as formed. The order of reaction in the presence of barium oxide was found to be zero throughout a run. The energy of activation of the inhibited reaction was 39 k. cal. per mole; that of the zero order rate in the presence of barium oxide, 26 k. cal. per mole. Measurements were also made of the heats of adsorption of the alcohol, 21 k. cal., and water, 13 k. cal., on bauxite. H. S. Taylor¹³ has pointed out the discrepancy which these figures seem to present. On the basis of current views, the true energy of activation of the inhibited reaction should be given by:

$$\begin{aligned}
 E_{\text{true}} &= E_{\text{apparent}} + \lambda_{\text{alc}} - \lambda_{\text{H}_2\text{O}} \\
 &= 39 + 21 - 13 \\
 &= 47 \text{ k. cal.}
 \end{aligned}$$

The zero order rate should yield at once the true energy of activation. It is observed to be 26 k. cal., not 47 k. cal. H. S. Taylor stresses the point that, if the heat of adsorption of water is only 13 k. cal. while that of the alcohol is 21 k. cal., water could not inhibit the adsorption of the alcohol. He suggests that a value nearer 30 k. cal. for the heat of adsorption of water would be more in keeping with the supposed poisoning by water, with the observed 26 k. cal. true energy of activation as also with the general chemisorption of water on oxide surfaces. It has never previously been pointed out that comparison of the zero order rate constant when barium oxide was present, with the zero order rate constant from the initial rates on bauxite alone, shows the former to be about ten times as great as the latter. Clearly, the Al_2O_3 and $\text{Al}_2\text{O}_3 \cdot \text{BaO}$ are not comparable catalysts, the latter being the more reactive, judged both by the actual rate and by the smaller energy of activation. There is, therefore, no proof that, on bauxite, the true order is zero; nor is it proved that the observed inhibition is attributable to poisoning by water. The fact remains that, on bauxite, the initial rate is zero order, but the course rate is first order. Naturally, too, the data are no justification for the relation between true and apparent energies of activation, a relation that is valid only for simultaneous adsorption of reactants or of products.

The generally accepted treatment of ammonia synthesis and decomposition on iron catalysts today, is that of Temkin and Pyzhev.¹⁵ The general formulation of these investigators gives a rate of decomposition proportional to $(\text{NH}_3^2/\text{H}_2^3)^\alpha$ and a synthesis rate proportional to $\text{N}_2(\text{H}_2^3/\text{NH}_3^2)^\alpha$. In the neighborhood of equilibrium they put $\alpha = 0.5$. Love and Emmett⁷ verified the relation for ammonia decomposition on a doubly promoted iron catalyst but found different kinetics for the singly promoted Fe-Al. Brill¹ reports the Temkin-Pyzhev relation to be an equally good approximation for ammonia synthesis on Fe, Fe-Al, and Fe-Al-K catalysts. It is, however, only an approximation, and the calculated constants using the Temkin-Pyzhev relation show a definite trend with increasing amounts of reaction. An invariant rate constant can be obtained, Brill shows, by choosing a suitable value for α , even though the value found necessary leads to an absurdity in the Temkin-Pyzhev derivation of the relation. Temkin and Kiperman have recently admitted that the Temkin-Pyzhev relation holds only in the immediate neighborhood of equilibrium, and they find, for extensive reaction, that α varies with the distance from equilibrium. In the case of synthesis, at a sufficient distance from equilibrium, $\alpha = 0$, and the rate is simply proportional to N_2 . It is interesting to recall the results of Seya,¹¹ for ammonia decomposition on reduced iron in a static system, which show a near-initial fall of the rate with progress of the reaction, described by the author as first order, followed by an extended range of constant rate, that is, zero order over the course of the reaction and terminated by a stage of falling rate only when the reactants become seriously exhausted. A similar result was reported by Kiyama,⁶ who used a doubly promoted iron catalyst.

It is apparent that the rates of ammonia synthesis and decomposition are, at present, expressed purely empirically, and that no sound general interpretation is available.

The three reactions considered are among the classical experiments which are frequently quoted as examples confirming the current view of the catalyst surface, a view involving essentially the Langmuir concept of adsorption. It is obvious that a reconsideration of this basic concept is not unwarranted. In the first place, according to Langmuir, adsorption is taken to represent an equilibrium between condensation and evaporation processes, and the surface a checkerboard with available sites for adsorption. The existence of an adsorption-desorption equilibrium cannot be denied in theory, but the practical problem remains whether the equilibrium corresponds to such an advanced stage of adsorption that the reverse desorption may be neglected, just as water dissociation is neglected in the hydrogen-oxygen reaction. It is certain that one of the first tests used to distinguish physical adsorption and chemisorption is the ease of desorption, it being recognized that the reversal of chemisorption requires, in general, a rather drastic change of conditions, such as high temperature, from those prevailing when chemisorption occurs; whereas a gas physically adsorbed is usually easily recovered. If desorption, that is, the evaporation process, is negligibly slow, the Langmuir isotherm as a relation between pressure and surface coverage has no meaning; that is, in the usual symbols, $\theta = 1$, coverage is effectively complete.

The checker-board picture of available sites on a surface, a site to be occupied by an atom or a molecule, has, as a necessary consequence, a momentary rate of adsorption proportional to the momentary surface concentration of available sites. Simple molecule-per-site occupancy demands first order kinetics. Taylor and Thon¹⁴ have shown that this is inconsistent with observations on rates of chemisorption and that, even when a chemisorption rate is initially first order, the rate is not maintained but falls exponentially as adsorption proceeds. This behavior has been attributed to a rapid production of sites by the gas on initial contact with the surface followed by a second order decay of these sites at a rate independent of the gas concentration. On this basis, the rate of slow chemisorption is governed solely by the availability of sites and the catalyst surface plays an active role as opposed to its passive role in the Langmuir picture.

In the application of the Langmuir view to a surface catalyzed reaction between two species, it has become the practice to assume that both species are simultaneously chemisorbed and the rate of the bimolecular reaction is taken as proportional to the product of the surface concentrations of the reactants. The catalyst is again assumed passive; constant in its activity during the reaction. The observed rates so interpreted give rise to an apparent order of reaction in each species. Thon and Taylor¹⁶ have analyzed the data for several typical bimolecular reactions, finding that a mere statement of an "order" is indefinite and may be seriously misleading unless a precise specification is made on whether the order refers to initial, or course of reaction, rates. There are many examples in the literature which indicate that the rate of a surface catalyzed reaction may decrease exponentially as the reaction proceeds, in a manner entirely analogous to the decay of sites in chemisorption. Thon and Taylor postulated a rate-determining role of the catalyst surface similar to that in chemisorption, the rate at each stage being proportional to the surface density

of active centers. An active center is interpreted as a dissociatively chemisorbed particle of one and only one reactant of the system. An opening step consists of reaction between the active center at the surface and a gaseous molecule of the other reactant, and is followed, in general, by further reactions with gaseous molecules to form final products. The active center density at the surface can be maintained over the course of the reaction through a chain mechanism.¹⁷ Alternatively, the center density may decay with progress of the reaction and the measured rate will decrease correspondingly. The rate of the opening step will involve the gas-phase concentration of one of the reactants. This reaction, which is responsible for its order, has led to the assumed requirement of its adsorption.

The presently accepted requirement for simultaneous adsorption and thus, presumably, for simultaneous activation of both reacting species stands in sharp contrast to the normal view for homogeneous reactions, particularly those involving radical chains. The primary basis of the Rice-Herzfeld or the Semenov-Hinshelwood schemes for radical reactions involves an initial production of radicals of one reactant only, not of both. Were both species to be activated, the resulting rate would be extremely slow. It seems illogical, therefore, to require it in heterogeneous reactions and not in homogeneous ones. The current view, requiring simultaneous adsorption of both reactants, especially if extensively dissociated, presupposes an obliging knowledge on the part of the surface atoms or radicals as to when and how they must desorb, verging on the absurd.

The current concept of the inhibition of reaction by a reactant or product is one of competitive occupation of the surface. For this reaction, the extent of adsorption of the poison must be greater than that of the other reactants. This result has actually been demonstrated unequivocally in remarkably few cases and, as pointed out here in the dehydration of isopropyl alcohol, is not supported by the experimental findings. In the quantitative application of this concept, for example, to the reaction of ethylene and hydrogen on nickel, it is the rule first to determine the fraction of the surface left bare by the more strongly adsorbed constituent. One would then suppose that the second reactant would avidly cover this remaining portion. The extent of coverage, however, is taken to be proportional only to the pressure of this reactant. One must conclude that some of the surface still remains bare, even though, in the absence of the poison, the second reactant at its pressure would occupy a greater proportion of the surface than the poison leaves bare. The significance of the alleged dense surface coverage is not obvious. It is well-known, too, that external catalyst poisons often inhibit markedly in amounts that are much too minute to correspond to a serious surface coverage.

This latter fact is, in part, one that has resulted in some modification of the simple Langmuir picture, by the postulation of a heterogeneity of the catalyst surface. The evidence for this assumption is based primarily on: (a) the decrease in the heat of adsorption with increased surface coverage; (b) a corresponding increase in the energy of activation of adsorption; and (c) inhibition by traces of poison.

Taylor and Thon have already shown that, in the course of chemisorption,

an interaction between sites, resulting in their decay, requires in general an exponential decrease in the heat of adsorption with increasing coverage.

A corresponding increase in the "energy of activation" of adsorption is also to be expected. This energy of activation is usually calculated by applying the Arrhenius equation to the times required for the adsorption of a given volume of gas at various temperatures; thus

$$-d \ln t/dT = E_{\text{ads}}/RT^2.$$

According to the Elovich equation

$$dq/dt = ae^{-\alpha q}$$

where q is the amount adsorbed in time t and a and α are constants.

Hence

$$t = (e^{\alpha q} - 1)/a\alpha$$

and

$$d \ln t/dT = \left(\frac{\alpha q e^{\alpha q}}{e^{\alpha q} - 1} - 1 \right) \frac{d \ln \alpha}{dT} - \frac{d \ln a}{dT}.$$

Taylor and Thon have shown that the variation with temperature of a , the initial rate of adsorption, may be expressed, $d \ln a/dT = E_0/RT^2$. On the other hand, α is either constant or decreases with increase in temperature. This variation would not be expected to be a simple one, since α is a quotient of the relative rate of site decay and the rate of chemisorption. Assuming, however, the relationship $-d \ln \alpha/dT = E_s/RT^2$, as an approximation, it follows that

$$E_{\text{ads}} = E_0 - E_s \left(\frac{\alpha q e^{\alpha q}}{e^{\alpha q} - 1} - 1 \right)$$

whence E_{ads} must increase as q increases since E_s is negative. There are very few data in the literature which cover a sufficient temperature range to justify a numerical test of this relation. The data of Burwell and H. S. Taylor² for the adsorption of hydrogen on chromic oxide which Taylor and Thon analyzed according to the Elovich equation¹⁴ offer a possible comparison. Burwell and Taylor state "if the curves be taken to represent activated adsorption alone, then between 427° and 405° K. the activation energy increases from zero to 18.5 k. cal. as the amount of hydrogen on the surface increases from 1.6 to 8.5 cc." These amounts of gas are adsorbed on 15 g. catalyst. It is not amiss to remark here that the zero value obtained for small coverage merely reflects an inability to derive the initial rate of adsorption from the observed data and not necessarily that the initial rate is temperature independent. Taylor and Thon find that a increases steadily with temperature corresponding to $E_0 = 12 \pm 2$ k. cal. The decrease of α with temperature was found to be much less uniform and yields a value for E_s of $-(15 \pm 5)$ k. cal. For 8.5 cc. adsorbed on 15 g., that is $q = 0.565$ cc. per g. at 1 atm. and 154° C., $\alpha = 2.7$ and the above relation gives a calculated E_{ads} of 24 ± 6 k. cal. as compared with the direct value of 18.5 k. cal. The Taylor-Thon interpretation requires that E_{ads} should increase from 12 ± 2 to 24 ± 6 k. cal. in the observed coverage range for which Burwell and Taylor found zero to 18.5 k. cal.

On the problem of inhibition, Thon and Taylor have shown¹⁶ that on the view that the inhibitor causes a widespread active center annihilation and does not inactivate by mere surface occupancy, not only can the relatively feeble inhibition by reactants and products be accounted for, but also the frequently catastrophic effects of traces of extraneous poisons.

On the three counts of varying heat of adsorption, varying energy of activation of adsorption, and inhibition it is apparent that heterogeneity of surface is not alone capable of rendering a satisfactory interpretation. It follows, therefore, that such evidence is not proof of that point of view. A chemical reaction, whether it occurs homogeneously or heterogeneously, is basically the same reaction and any apparent difference can lie only in the specificity of the course of the reaction owing to the involvement of the surface. The chain interpretation of heterogeneous catalysis by Thon and Taylor offers the parallel with homogeneous reaction.

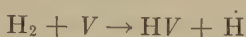
A possible alternative reinterpretation of the schemes recently advanced by Taylor and Thon for the kinetics of chemisorption and of heterogeneously catalyzed reactions may be of interest in the light of ideas expressed in a review article by N. N. Semenov,¹⁰ which appeared after the first presentation of the content of the first communication,¹⁴ and was received and came to Doctor Thon's attention only after the second communication¹⁶ had been presented for publication. It is opportune to present briefly Semenov's ideas on chemisorption and catalysis and to examine their implications.

Semenov assumes the original presence, on the catalyst surface, of a certain number of free valences V which can radicalize an inert species striking from the gas phase. On the other hand, if a free radical strikes an inactive point, S of the surface, it will give rise to a surface dissociation $S \rightarrow V + V$; one V will saturate the radical, and the second will remain available for reaction with an inert species. Semenov's example is $V + C_2H_4 \rightarrow VCH_2\dot{C}H_2$; $VCH_2\dot{C}H_2 + S \rightarrow VCH_2CH_2V + V$. Applying this idea to the chemisorption of hydrogen, one can write $V + H_2 \rightarrow VH + \dot{H}$; $\dot{H} + S \rightarrow HV + V$, *etc.* This provides a chain mechanism for chemisorption, which could be the way in which chemisorption sites are produced through contact with the gas as concluded by Taylor and Thon.

Because of the resonance of bonds between neighboring atoms on a surface, it is not unlikely that this chain initiation of sites could become almost explosive by branching chains and account for the very rapid near-initial site incubation period to which Taylor and Thon have attributed the initial massive adsorption so frequently observed in chemisorption studies. On the other hand, the bimolecular decay of the site density n , previously postulated, can be interpreted as the second order rate of a site recombination reaction $V + V \rightarrow S$.

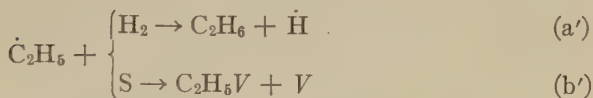
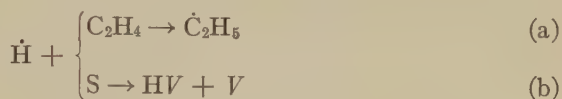
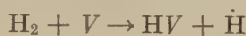
For catalytic reactions, Semenov's idea amounts to attributing to the catalyst surface solely the role of radical generator and chain terminator, the acts giving rise to the reaction product proper, taking place between a radical and a physically adsorbed inert molecule. The specific necessity of physical adsorption is questionable, as in this scheme the reacting radicals are not those chemisorbed but, on the contrary, those liberated. In application to the

$\text{H}_2 + \text{D}_2$ exchange, Semenov would write



and the analogous reactions starting with D_2 . In either case, *a* or *b*, the chain is carried forward; in the event *a*, it is through $\dot{\text{D}}$, in *b* through V . In this view, only $\dot{\text{H}}(\dot{\text{D}})$ and V are reactive, but $\text{HV}(\text{DV})$ are not.

For the hydrogenation of C_2H_4 with the earlier assumption¹⁶ of H_2 as the sole center-making reactant,*



The common ground of the Thon-Taylor formulation and the above schemes based on Semenov is the possibility of a chain character of surface-catalyzed reactions, and the definite rejection of reactions between two chemisorbed species. The difference is that in the Thon-Taylor scheme the act giving rise to the product takes place between a chemisorbed species and a gaseous reactant, whereas in Semenov's the product, even in heterogeneous catalysis, is formed in a reaction between a free radical and an inert molecule, exactly as in homogeneous chain reactions. This view is essentially tantamount to the "hetero-homogeneous" concept long defended by Polyakov.⁸

On the basis of Semenov's mechanism, instances of the applicability of the Elovich exponential rate law can be interpreted as follows. The initiation step of the chain evidently requires an active surface site V . The number of chains initiated per unit time will therefore be proportional to the momentary site density, n . If the sites are subject to bimolecular decay at a rate proportional to n^2 , the rate of initiation, and consequently, the course of the reaction, will follow, as was previously shown, a rate law $dq/dt = ae^{-\alpha q}$; the amount reacted, q , and the amount chemisorbed, are evidently proportional to each other, since each initiation involves also chemisorption.

A decision between these two alternative interpretations of the nature of the chain carriers in surface-catalyzed reactions does not appear obvious for the present.

* Semenov's way of starting with $V\text{CH}_2\dot{\text{C}}\text{H}_2$ does not seem to be consistent with the kinetic fact that the initial rate is nearly first order in H_2 , but zero or slightly negative in C_2H_4 .

References

1. BRILL, R. 1951. *J. Chem. Phys.* **19**: 1047.
2. BURWELL, R. L. & H. S. TAYLOR. 1936. *J. Am. Chem. Soc.* **58**: 697.
3. DOHSE, H. & W. KÄLBERER. 1929. *Z. physik. Chem. (B)* **5**: 131.
4. HAUFFE, K., R. GLANG & H. J. ENGEL. 1952. *Z. physik. Chem.* **201**: 223.
5. HINSHELWOOD, C. N. & C. R. PRITCHARD. 1925. *J. Chem. Soc.* **127**: 327.
6. KIYAMA, R. 1939. *Rev. Phys. Chem. Japan.* **13**: 125.
7. LOVE, K. S. & P. H. EMMETT. 1941. *J. Am. Chem. Soc.* **63**: 3297.
8. M. V. POLYAKOV. 1948. *Uspekhi Khim.* **17**: 351.
9. SCHWAB, G. M., R. STAEGE & H. H. v. BAUMBACH. 1933. *Z. physik. Chem. (B)* **21**: 65.
10. SEMENOV, N. N. 1951. *Uspekhi Khim.* **20**: 673.
11. SEYA, K. 1939. *Rev. Phys. Chem. Japan.* **13**: 137.
12. STEACIE, E. W. R. & J. W. McCUBBIN. 1934. *J. Chem. Phys.* **2**: 585.
13. TAYLOR, H. S. 1931. *J. Am. Chem. Soc.* **53**: 595.
14. TAYLOR, H. A. & N. THON. 1952. *J. Am. Chem. Soc.* **74**: 4169.
15. TEMKIN, M. I. & V. PYZHEV. 1940. *Acta Physicochim. U. R. S. S.* **12**: 327.
16. THON, N. & H. A. TAYLOR. 1953. *J. Am. Chem. Soc.* **75**: 2747.
17. WAGNER, C. & K. HAUFFE. 1939. *Z. Elektrochem.* **45**: 409.

THE EFFECT OF MONOLAYERS ON THE RATE OF EVAPORATION OF WATER

By Robert J. Archer and Victor K. LaMer

Columbia University, New York, N. Y.

Summary

The development of an experimental technique for measuring the influence of surface films on the rate of evaporation of water is described. The effect is measured by the rate of absorption of water vapor by a solid desiccant suspended above the water surface. A mathematical analysis relating the rate of evaporation measured above the water surface to the actual rate at the surface is given. The effect on the rate is expressed as the specific resistance of the monolayer to evaporation which is inversely proportional to the absolute rate. Measurements of the specific resistances of three long chain fatty acids are described. The plot of specific resistance as a function of surface pressure is interpreted in terms of the occlusion of impurities in the monolayer. When the sources of these impurities are eliminated, the specific resistance is independent of the surface pressure but depends on the chain length of the film molecules. A theory for the resistance of monolayers is proposed in terms of an energy barrier attributable to the film. The magnitude of this barrier is calculated theoretically and agrees reasonably well with the experimentally determined value. The data and ideas of previous workers are discussed.

Technique and Theory of Measurements

The technique used is a development of the method of Langmuir and Schaefer,¹ who measure the rate of absorption of water vapor by a granular solid desiccant supported in a container above the water surface. A quantity defined as the evaporation resistance of the film is proportional to the difference in the reciprocals of the rates of absorption with and without a surface film.

The measurements are made in the trough of a surface balance, so that the resistance to evaporation can be determined as a function of the film pressure and molecular area of the monolayer. FIGURE 1 is a cross sectional sketch of the apparatus. The desiccant container *C* is a cylindrical leucite box. The bottom is open except for a membrane *M* which retains the desiccant but is permeable to water vapor. The thermometer *T'*, inserted through the rubber stopper *S*, measures the temperature of the desiccant. An airtight leucite lid fits the bottom face when the container is not in use. In making a measurement, the container fits into the leucite ring *R* and is supported by the extended rim *L*. The supporting ring is inserted in a circular hole in the square leucite plate *P* which rests on the edges of the trough *B* of the surface balance, thus holding the bottom face of the container parallel to and about 2 mm. from the water surface. A thermometer is placed just beneath the water surface.

To determine the rate of absorption, the container, filled with desiccant, is weighed; the lid is removed quickly and the container put into position above the water surface for a definite length of time; it is then removed, the lid

quickly replaced and the container reweighed. The increase in mass divided by the time in position gives the rate of absorption.

A number of materials were tried for the desiccant container membrane, including fine copper gauze and various cloths and papers. Most were impractical for one reason or another. Copper gauze and some cloths allowed sifting of the desiccant through the membrane. Paper was generally insufficiently permeable to water vapor, and both paper and most cloths became moist with consequent peculiarities in the rate of absorption. With most materials, after one or two absorption runs, a thin layer of desiccant adhered to the membrane and could not be shaken loose. Since mixing of the desiccant before each run to insure homogeneity was necessary to insure reproducibility, this complication had to be avoided. The material best suited for the membrane is the heavy silk-like waterproofed cloth used to make umbrellas. It is permeable to water vapor, allows little sifting and does not become moist or form an adhering layer of desiccant.

The choice of desiccant is limited by the following factors: it should have a constant aqueous tension for a considerable range of absorbed water; it should not deliquesce; and it should have a rate, an enthalpy, and a specific heat of absorption such that the temperature changes during absorption will not be excessive. The substances which meet these requirements best are lithium chloride and calcium chloride. Anhydrous lithium chloride is the better choice since it produces a greater rate of absorption and exhibits less caking and adherence to the membrane.

The analysis of the transport of water from the water surface to the desiccant surface involves evaporation from the water surface, diffusion through the separating air column and the membrane followed by condensation or absorption at the desiccant surface. The symbols used in the analysis are defined in TABLE 1.

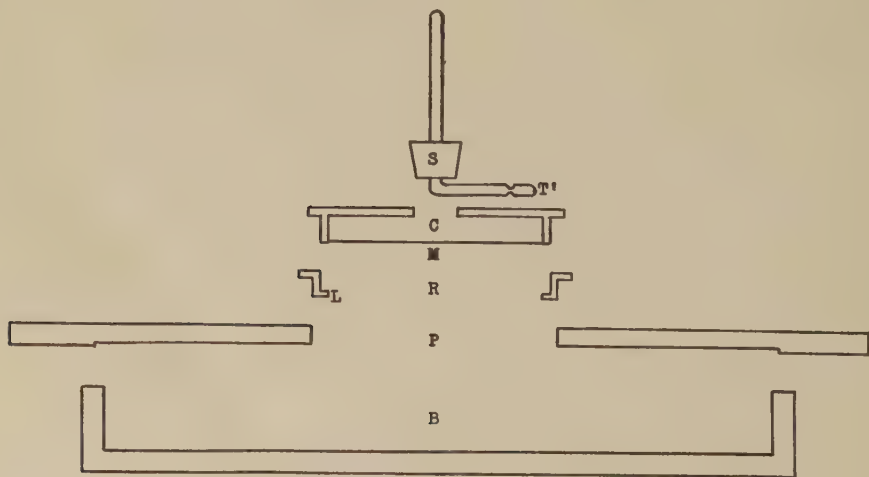


FIGURE 1. The desiccant container and supporting apparatus.

TABLE 1
GLOSSARY OF SYMBOLS

M	= mass of water vapor absorbed by the desiccant.
t	= time of absorption.
A	= area of the desiccant surface.
b	= distance between the water and membrane surfaces.
b'	= thickness of the membrane.
T	= temperature of the water surface.
T'	= temperature of the desiccant surface.
m	= mass of a water molecule.
Q	= $(kT/2\pi m)^{1/2}$.
Q'	= $(kT'/2\pi m)^{1/2}$.
w	= concentration of water vapor in equilibrium with liquid water at temperature T .
w_0	= concentration of water vapor in equilibrium with the desiccant at temperature T' .
w'	= concentration of water vapor at water surface during a measurement.
w_0'	= concentration of water vapor at desiccant surface during a measurement.
w_m	= concentration of water vapor at membrane surface during a measurement.
D	= diffusion coefficient for water vapor in air.
D'	= diffusion coefficient for water vapor in membrane.
α	= fraction of water molecules impinging on water surface which condense.
β	= fraction of water molecules impinging on desiccant surface which are absorbed.

The net rate of evaporation from the water surface is the difference between the rate of evaporation and the rate of condensation. If these rates are described as the fraction of the vapor molecules impinging on the surface which condense, then the net rate of evaporation is given by

$$M/t = A\alpha Q(w - w'). \quad (1)$$

The rate of diffusion from the water surface to the membrane is

$$M/t = DA(w' - w_m)/b. \quad (2)$$

The rate of diffusion through the membrane is given by

$$M/t = D'A(w_m - w_0')/b'. \quad (3)$$

The rate of absorption at the desiccant surface is

$$M/t = A\beta Q'(w_0' - w_0). \quad (4)$$

If these four equations are combined to eliminate w_0' , (w_m) , and w' , the result is

$$M/(At) = (w - w_0)/(b/D + b'/D' + 1/\beta Q' + 1/\alpha Q). \quad (5)$$

This is the basic equation relating the rate of absorption by the desiccant to the actual rate of evaporation at the water surface in terms of the condensation coefficient α .

EQUATION 5 suggests an analogy to Ohm's law in the form

$$\text{Rate} = (\text{Driving force})/(\text{Resistance})$$

thus

$$M/t = (w - w_0)/R,$$

where R is the sum of the resistances of the air column, the membrane, the

desiccant surface and the water surface acting in series. If the specific resistance is defined as

$$r = RA,$$

the four terms in the denominator of the right-hand side of EQUATION 5 are, respectively, the specific resistances of the air column, the membrane, the desiccant surface, and the water surface.

The effect of surface films on the evaporation rate can now be expressed in terms of the specific evaporation resistance of the film, which is related to the condensation coefficient of the film-covered surface by the expression

$$r = 1/\alpha Q. \quad (6)$$

Henceforth, r will refer only to the specific resistance of the water surface or the monolayer with the dimensions of sec./cm.

If two determinations are made of the rate of absorption, one with a monolayer on the surface and one without, the specific resistance of the monolayer is calculated from the expression

$$r = [A(w - w_0)t/M]_{\text{film}} - [A(w - w_0)t/M]_{\text{no film}}, \quad (7)$$

which follows from EQUATION 5 when b , T , and T' are the same for the two determinations and when the specific resistance of the monolayer is much greater than that for a clean water surface, which is always the case.

In deriving EQUATION 5 it was assumed that the various parts of the system remained at constant temperature during a measurement of M . Actually, the temperature of the water surface decreases with evaporation; the desiccant temperature increases, and a varying temperature gradient exists in the air column. It will be assumed that the average temperatures of the water surface, the air column and desiccant surface can be used to determine w , w_0 , and D in EQUATION 5. These average temperatures are estimated as follows:

In order to minimize temperature variations in the water surface, water from a thermostat is pumped through a pyrex glass serpentine tube in the trough. At some point below the water surface, say at the mid-point of the serpentine tube, the temperature remains constant. This temperature is measured by a thermometer on the bottom of the trough. The temperature, T_t , recorded by the thermometer whose bulb is just below the water surface, is assumed to be the mean temperature of the gradient between the plane at constant temperature T_c and the water surface. Therefore, if T_s is the surface temperature,

$$T_s = 2T_t - T_c.$$

During absorption, T_t drops rapidly to an approximately constant value (the temperature change is usually 0.1 to 0.5° C.) so that the effective average temperature of the water surface during a run is given by this expression when the value of T_t is recorded just before the desiccant container is removed. If T_t does not remain sufficiently constant during a run, the average of the readings of the thermometer taken at one minute intervals is used to determine $\bar{T}_s$.

During a run, a temperature gradient is set up between the lower and upper

desiccant surfaces. The temperature of the upper surface is assumed to be that of the thermometer reading at that point. If the contents of the container are thoroughly mixed and the thermometer bulb submerged in the desiccant, the thermometer approximates the average temperature of the gradient. Therefore, if the temperatures of the lower surface, the upper surface and the average temperature are designated by T_s' , T_c' and T_t' ,

$$T_s' = 2T_t' - T_c'.$$

For absorption runs of the duration used, the average temperature and the temperature of the upper surface increase linearly with time, and presumably the lower surface temperature does also. The average temperature of the lower surface will therefore be the average of the two values for T_s' calculated for values of T_t' and T_c' measured immediately before and after the absorption run.

Since a temperature gradient exists in the air column with the average temperatures of the opposite faces equal to $\bar{T}_s'$ and $\bar{T}_s$, the value taken for the air column is

$$\bar{T}_{\text{air}} = (\bar{T}_s + \bar{T}_s')/2.$$

There will be an additional contribution to M from water vapor in the surrounding air. It is necessary to subtract this contribution from the total mass absorbed to obtain the correct value for M for use in the equations developed above. The correction is determined under the standard conditions by covering the surface with a thin sheet of aluminum foil instead of a monolayer. The rate of absorption decreases with time, being rapid at first and becoming much slower with the exhaustion of the water vapor in the surrounding air. Under typical experimental conditions with the aluminum foil on the surface, 0.0085 g. of water is absorbed in three minutes, 0.0095 g. in four minutes and 0.0125 g. in eight minutes. For comparison, the amounts of water absorbed without the aluminum foil for times of three, four, and eight minutes are 0.250 g., 0.333 g., and 0.666 g. This correction also includes the water vapor absorbed while the desiccant container is being put in place above the water surface, before timing is begun and the amount absorbed while the container is removed from the supporting ring and the lid is put in place, after timing is stopped.

If a series of determinations of M as a function of t are made with a clean water surface and constant b and D , the following relation holds:

$$M = \text{constant} (w - w_0)t.$$

Therefore, if the correction is valid, a plot of M versus $(w - w_0)t$ will be a straight line passing through the origin. This requirement was confirmed experimentally.

The validity of EQUATION 5 and of the approximations proposed can be established by measuring one of the specific resistances as a function of $A(w - w_0)t/M$ while holding the other resistances constant. This measurement is made by plotting $DA(w - w_0)t/M$ as a function of the distance between the desiccant surface and the water surface. This distance can be varied by introducing supporting rings R of FIGURE 1 with varying heights of the rim

L . Seven rings were constructed with L varying from 0 to 0.6 cm. in increments of 0.1 cm. The position of the water surface is held constant to within 0.0015 cm. by means of a hydrophilic point of glass submerged in the trough. When the height of the water surface decreases, the glass point protrudes through the surface. Water is then added until the point just disappears below the surface.

If B is the height of the rim L and b_0 is the constant distance between the water surface and the bottom surface of the supporting plate, EQUATION 5 can be rewritten

$$\begin{aligned} B &= DA(w - w_0)t/M - (b_0 + Db'/D' + D/\beta Q' + D/\alpha Q) \\ &= DA(w - w_0)t/M - \text{constant.} \end{aligned} \quad (8)$$

So that the second term on the right-hand side of EQUATION 8 will be constant, D is held constant by cooling the desiccant container, when necessary, to such an initial temperature that the average temperature of the air column is constant for a series of determinations.

FIGURE 2 shows a typical experimental plot for testing EQUATION 8. The slope is unity within 1.5 per cent, a variation to be anticipated from the uncertainties in the values of the quantities, and the validity of the technique for measuring specific resistance is established.

Of the quantities on the right-hand side of EQUATION 8, A , t and M are measured directly. The diffusion coefficient for water vapor in air is calculated from the expression

$$D = (0.220/p)(T/273)^{1.75} \text{ cm.}^2/\text{sec.}$$

from the International Critical Tables. In this equation, p is the atmospheric pressure in atmospheres, and T is the temperature in $^{\circ}\text{K}$.

In these experiments, the desiccant surface became sufficiently moist to behave as a film of saturated solution. A value for w_0 corresponding to the aqueous tension of a saturated lithium chloride solution rather than the ten-fold smaller value corresponding to the aqueous tension of the solid monohydrate satisfied EQUATION 8. Similar results were obtained using calcium chloride, *i.e.*, the aqueous tension of the saturated solution fitted the data, whereas values for the solid hydrates did not. The value for w_0 of lithium chloride is taken from graphical data published by Bichowsky² and the *Heating, Ventilating and Air Conditioning Guide* for 1941.

Experimental Results

The specific resistances of monolayers of stearic, nonadecanoic, and arachidic acids have been studied. They will be referred to as the C_{18} , C_{19} , and C_{20} acids. Melting point determinations indicate initial preparations of high purity, which were further purified by recrystallizing from twice-distilled petroleum ether. The resistance is measured as a function of surface pressure using a standard surface balance arrangement of the vertical-pull type described by Harkins.³ The water and organic solvents were twice distilled from an all Pyrex glass apparatus.

FIGURE 3 is a typical pressure-molecular area isotherm for monolayers of long

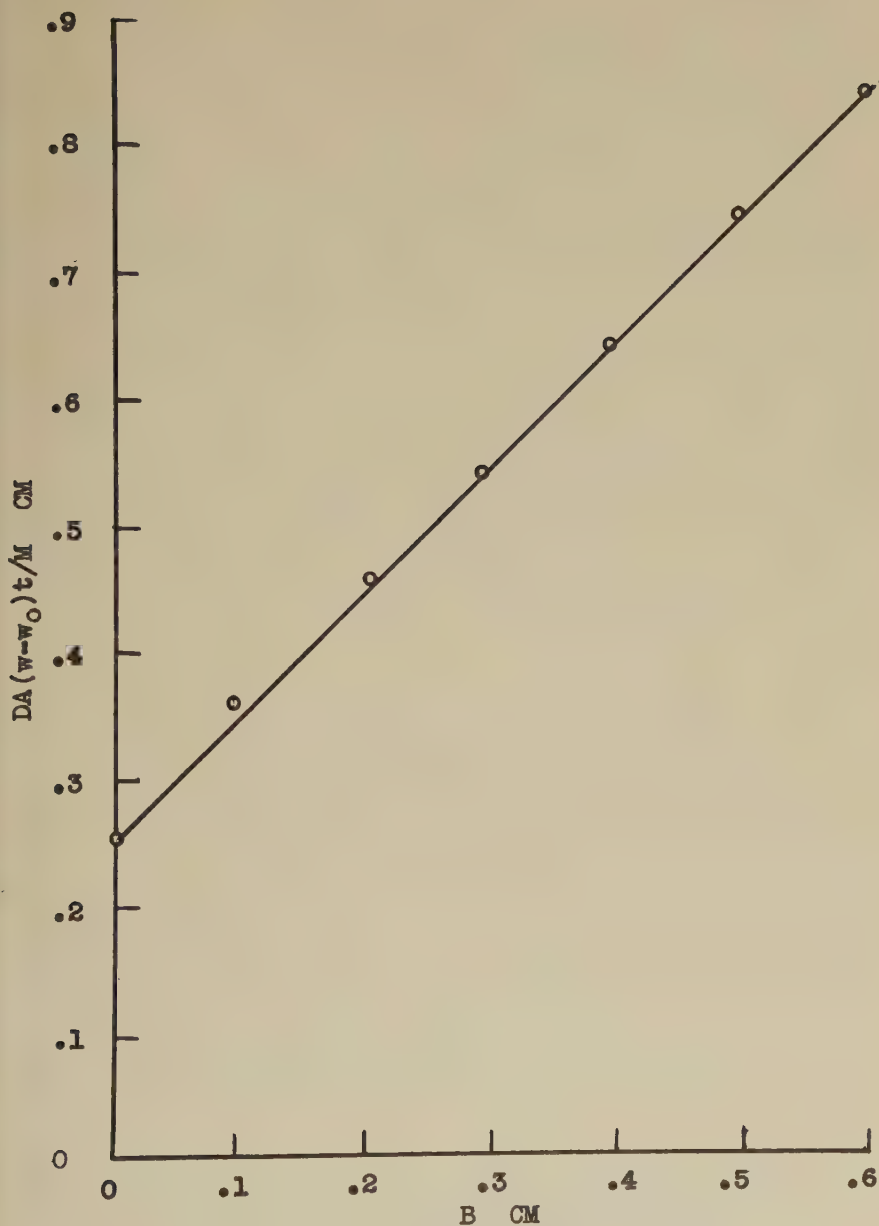


FIGURE 2. The resistance of the air column and validation of technique.

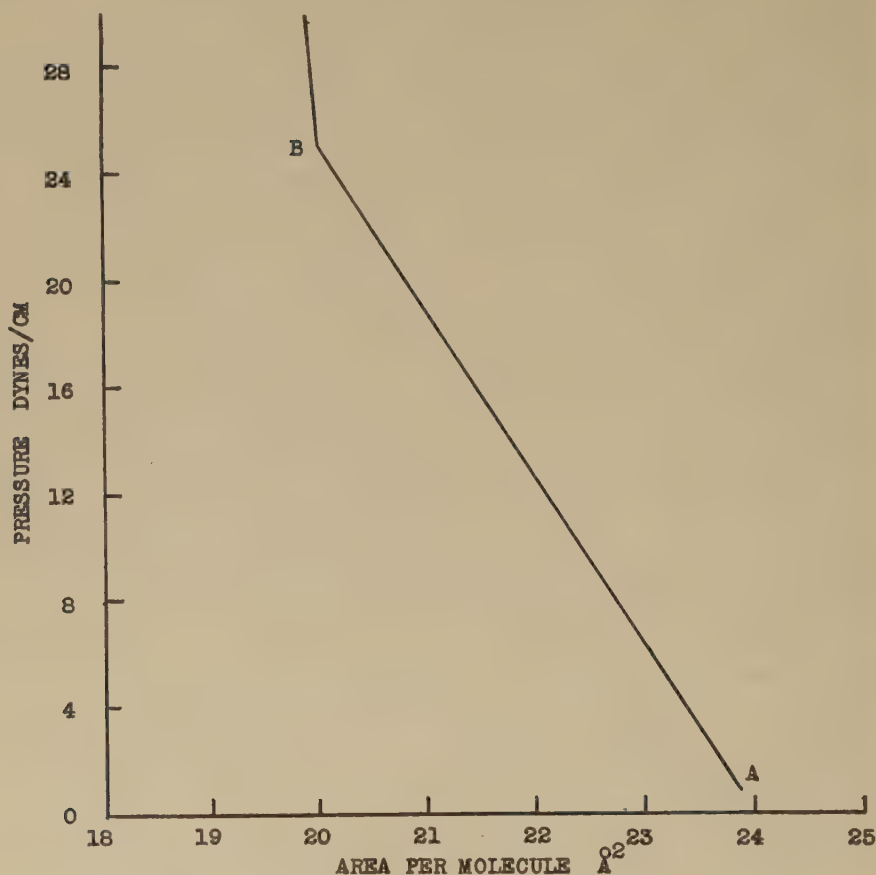


FIGURE 3. Pressure-molecular area isotherm for long chain fatty acids.

chain acids at 25°C . The lower arm of the plot represents the liquid condensed phase of the film and the upper arm the solid phase. The first detectable change in surface tension occurs at a molecular area of about $24\text{ \AA}^2$, from which point the increase in pressure with compression is linear. At about $20\text{ \AA}^2$ per molecule, the phase change occurs with a sudden decrease in the compressibility of the film. The differences in this isotherm for acids of different chain lengths are minor and need not be considered.

The measurements of the resistances of these films as a function of pressure in the solid phase are ambiguous, since a slow collapse of the monolayer begins as soon as the transition point is passed. The pressure can decrease as much as 5 dynes/cm. during the time a resistance measurement is made. For this reason, the quantitative measurements of FIGURE 5 were not extended into the solid phase.

The characteristics of the plot of specific resistance as a function of film pres-

sure are very sensitive to variations in the technique used to spread the film. Although the source of this sensitivity is now believed to be known and eliminated, a discussion of the effects of changes in technique on the p - r plot is in order since an understanding of these effects clarifies the nature of the resistance of monolayers and the significance of the results reported by previous workers.

The factors of importance in determining the shape of the p - r curve and the magnitude of r at a given pressure are: (1) the choice of solvent used for the spreading solution; (2) the concentration of the spreading solution; (3) the pressure or molecular area of the film immediately after spreading the film and before compression is begun; (4) the cleanliness of the water surface immediately prior to spreading the film; and (5) the paraffin used to coat the inside surfaces and edges of the trough.

The metamorphosis of the p - r plot for the C_{19} acid on distilled water at 25°C ., for specific alterations in technique from the first measurements to what are believed to be the correct values, is shown in FIGURE 4. Each set of data represents typical reproducible results for a fixed technique of spreading. For curves I through IV, the trough was a chromium-plated brass tray coated with paraffin on the inside surfaces and upper edges. It was cleaned by flushing with hot water and the paraffin was renewed periodically. The initial pressure referred to in each case below is the surface pressure of the film before compression. In general, a pressure of zero means that the initial total area of the film must be decreased by at least one seventh before any change in surface tension is detectable.

Curve I was obtained using benzene as the spreading solvent for a solution 0.0035 molar in C_{19} acid and zero initial pressure. Curve II results from substituting petroleum ether as the solvent and using the same concentration and initial pressure as for curve I. Curve III was obtained by doubling the concentration of the acid in the petroleum ether solution. If the concentration of the petroleum ether solution is again doubled (now 0.014 molar), an initial pressure of about 5-10 dynes/cm. is used and special care is taken to sweep the water surface clean as soon as possible before applying the spreading solution, curve IV is obtained. Curve IV can also be obtained by using ethyl ether or pentane as the spreading solvent. Finally, using the trough described, no further variation in technique, such as increasing the concentration of the spreading solution, changes the characteristics of curve IV.

If the concentration of acid in the benzene spreading solution is increased, there is a concomitant displacement of the p - r curve upward on the r axis, but it always lags behind the petroleum ether result, and in no case is curve IV obtained.

The following is a summary of these results. Benzene spreading solutions yield consistently low values for specific resistance, and the shape of the p r curve is approximately exponential. Petroleum ether solutions give larger values for r , and the p r curves are nearly linear. Increasing the concentration of the acid in the petroleum ether solutions increases the magnitude of r until a limiting concentration of about 0.014 molar is reached, beyond which no further concentration effect is obtained. Finally, if special care is taken to

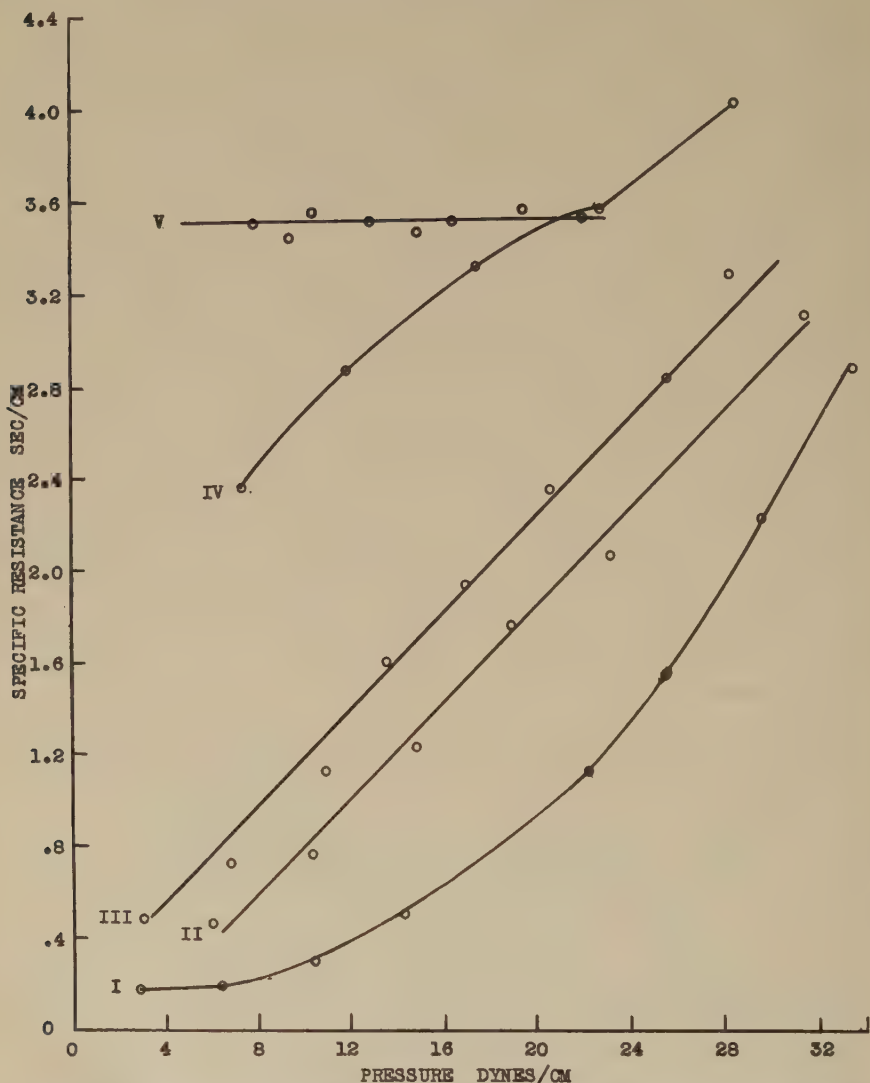


FIGURE 4. The effect of spreading technique variations on specific resistance as a function of surface pressure for the C_{19} acid on distilled water at $25^{\circ}C$.

sweep the surface thoroughly clean immediately before applying the spreading solution and the initial pressure is greater than zero, a maximum result is obtained for the paraffined trough.

It is believed that these variations in the p - r plot are attributable to impurities, principally solvent molecules, occluded in the monolayer as a result of the spreading technique and that the increase of r with p is attributable principally to the expulsion of these occlusions on compression of the film. As will be

shown below, the resistance of monolayers is related to the energy required for an escaping water molecule to form a hole in the surface film by displacing one of the film molecules. The magnitude of this energy depends on the potential energy of interaction of a film molecule with all of its neighbors in the monolayer. Small foreign molecules having relatively meager interactions with the surrounding molecules constitute permanent holes in the film or at least sites of small resistance. Since the total resistance of the monolayer is the resistance of these sites acting in parallel with sites occupied by acid molecules, a very small concentration of occluded foreign molecules could produce a large decrease in the net resistance. The following calculation will illustrate this point. Assuming that the foreign molecules constitute free sites, the specific resistance of these sites is that of a clean water surface, which is given by

$$r_0 = 1/\alpha_0 Q = 1.9 \times 10^{-3} \text{ sec./cm.}$$

if, after Alty,⁴ the value of α_0 (the condensation coefficient for a clean water surface) is taken as 0.036.

For the C_{19} acid, as will be seen below,

$$r_{C_{19}} = 3.5 \text{ sec./cm.},$$

and the net resistance for small concentrations of occluded molecules will be given by

$$1/r = n/r_{C_{19}} + n_0/r_0$$

where n and n_0 signify the mole fractions of the acid and the foreign molecules. From this relation the specific resistances of monolayers of the C_{19} acid with 1 per cent, 0.1 per cent, and 0.01 per cent concentrations of occluded foreign molecules are, respectively, 0.16, 1.2, and 2.9 sec./cm. These values, compared to 3.5 for an uncontaminated film, show the large decrease in resistance that accompanies occlusion of foreign molecules in the film.

Since the specific resistances of benzene molecules and of the short chain molecules of petroleum ether probably have about the same magnitude, more solvent molecules must be occluded when benzene is the spreading solvent than when petroleum ether is used to explain the results of FIGURE 4. The relative degree of occlusion will depend on the ability of the occluded molecules to interact both with the neighboring long chain acid molecules and with the water surface. That this interaction would be greater for benzene molecules follows from the fact that long chain acids are more soluble in benzene than in petroleum ether and that benzene itself will form a monolayer on water with an equilibrium film pressure of about 10 dynes/cm., indicating considerable interaction between benzene molecules and the water surface. A further factor favoring the occlusion and retention of benzene molecules is the lower volatility of benzene compared to petroleum ether.

Besides the occlusion of solvent molecules, three other sources of occluded impurities are present in the system: namely, the substrate water, the air above the water surface and the paraffined surfaces of the trough.

The effect of impurities which settle on the water surface from the air and which are absorbed at the surface from the underlying water before the monolayer is spread will be greatly decreased by thoroughly cleaning the water surface as soon as possible before applying the spreading solution. This experimental precaution does result in higher values for τ . Absorption of impurities at the surface after the monolayer is formed is minimized by maintaining the film at relatively high pressures since the rate of absorption is known to depend on the compression of the film. This explains the better values of τ when initial pressures above zero are used.

A likely source of occludable impurities is the large area of paraffined surface of the metal trough. Therefore, an all glass trough of the type described by Stenhagen and others⁵ was constructed. It was made by sandblasting a cavity 1.5 cm. deep, 65 cm. long, and 14 cm. wide in a solid block of Pyrex glass 2.5 cm. thick. It is unnecessary to paraffin the inner surfaces of this trough, and it can be cleaned easily with chromic acid cleaning solution. It is, of course, still necessary to paraffin the upper edges so that they will be hydrophobic, but the amount of paraffin used was minimized by lightly grinding these surfaces and by applying the paraffin by rubbing a stick of the solid over the ground surfaces and then rubbing with a clean towel until only a thin film was retained, which could be removed only with a solvent. In this way, the edges were rendered hydrophobic with a minimum amount of paraffin.

When resistance measurements were made using this trough and the technique of curve IV, curve V (FIGURE 4) was obtained. This curve is effectively independent of the pressure of the film within the range measured, indicating that the pressure dependent "tail" of curve IV was due to absorption of impurities from the paraffined surfaces.

It is assumed that measurements employing the technique of curve V give correct values for the specific resistance of fatty acid monolayers, and that this technique is essential for the formation of such monolayers if they are to be free of impurities.

Employing the technique of curve V, specific resistance as a function of pressure was measured for the three long chain fatty acids C_{18} , C_{19} and C_{20} at 25° C. on double-distilled water. FIGURE 5 shows the results of these measurements. As with the C_{19} measurement just described, all three curves are independent of surface pressure within the range investigated. Although occasionally a low pressure point falls below the horizontal line, no point is ever significantly above this line when the film is in the liquid condensed phase. In the range from 7 to 25 dynes/cm. any pressure dependence of specific resistance is small enough to be obscured by the uncertainty in the measurements and by effects attributable to slight contamination.

An increment of one $-CH_2-$ group in the fatty acid chain has a large effect on the resistance of the monolayer. The resistance of the C_{19} film is 1.3 sec./cm. above the C_{18} value, and the C_{20} result is 1.95 sec./cm. greater than the C_{19} resistance.

Extension of the measurements of τ beyond 25 dynes/cm., *i.e.*, into the solid phase, indicates that, in this range, specific resistance increases with pressure, but the measurements are not sufficient to warrant description at the present

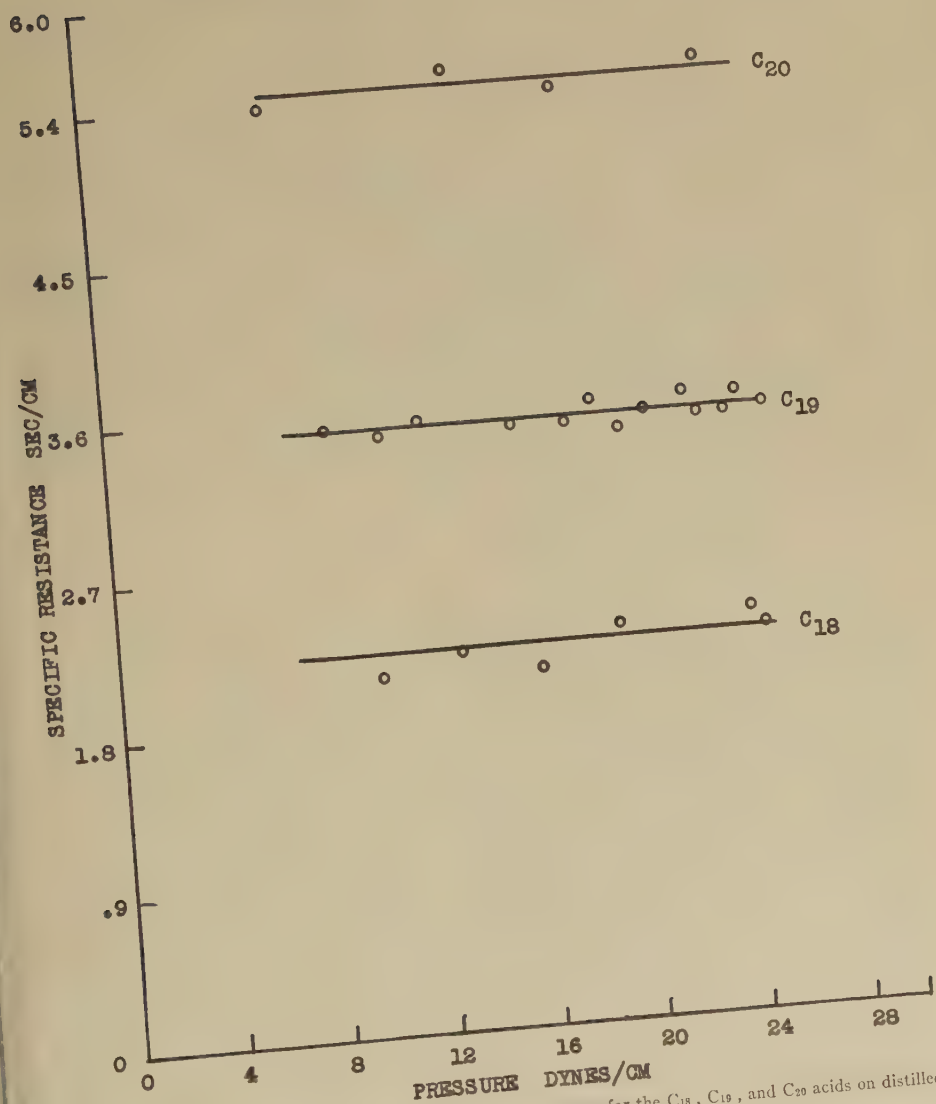


FIGURE 5. Specific resistance as a function of surface pressure for the C₁₈, C₁₉, and C₂₀ acids on distilled water at 25° C.

time. Significant data for the solid phase are difficult, if not impossible, to obtain because of the collapse phenomenon mentioned above.

The results of FIGURE 5 in terms of the actual effect of the films on the evaporation rate can be expressed as the ratio of the evaporation rate with and without a monolayer, namely, r_0/r . These ratios, times 10^4 , are, respectively, 8.3, 5.4, and 3.5 for the C₁₈, C₁₉ and C₂₀ acid monolayers. A long chain

fatty acid monolayer decreases the rate of evaporation of water by a factor of about 10^4 .

Theory of the Resistance of Monolayers to Evaporation

The resistance to evaporation is attributed to an energy barrier resident in the film. The magnitude of this barrier is determined by the energy required by a condensing or evaporating water molecule to penetrate the monolayer. Since the condensation coefficient α is defined as the fraction of impinging water molecules which penetrate the film, it is just this fraction of water vapor molecules which have an energy available for penetration equal to or greater than the energy barrier. It is not possible to predict a distribution function which will describe α , as there is no *a priori* way of determining which degrees of freedom are available for the work of penetration. Acceptable agreement between theory and experiment is obtained if α is expressed by the classical Maxwellian distribution which describes the fraction of the vapor molecules which have an energy in three degrees of freedom equal to or greater than any arbitrarily selected value. This description indicates that three degrees of freedom are available for the work of penetration, which are probably those associated with translational motion.

Thus, specific resistance is given by

$$r = (\pi^2 m / 2E)^{1/2} \exp(E/RT) [1 + 1/(2E/RT) - 1/(2E/RT)^2 + \dots]^{-1}, \quad (9)$$

where E is the energy barrier to evaporation, and r can be calculated theoretically by postulating a mechanism for penetration and a structure of the monolayer from which the energy barrier can be estimated. It is assumed that, over limited areas, the molecules in a monolayer are in a two-dimensional close-packed hexagonal array and that a water molecule in penetrating the film must form a hole in the monolayer equivalent to the area occupied by one film molecule.

Consider a small area of the hexagonal array of molecules. In the initial state all sites are occupied. After the hole is formed, there is one free site in the small area, but the mechanism of displacement is such that the total number of molecules in the area is unchanged. Consequently, the intermolecular separation has been decreased, but for a sufficiently large area this change is negligible. The energy change resulting from the formation of the free site is expressed in terms of the potential energy of interaction between two molecules in the monolayer designated by w . In the initial state, the total interaction energy is given by the summation, over all of the molecules, of the energy of interaction between two molecules, w_{pq} , that is,

$$\sum_p \sum_q w_{pq}.$$

In the final state, the total interaction energy is this summation less the interaction between all of the molecules in the area and the molecule formerly occupying the free site, which is

$$\sum_q \sum_p w_{pq} - \sum_p w_{ap}.$$

Consequently, the energy change is just the interaction between one molecule and all of its neighbors in the monolayer, and is given by

$$W = \sum_p w_{ap}.$$

Hence, W is the magnitude of the energy barrier E in EQUATION 9.

Joly⁶ has developed a method for estimating the interaction energy between long chain molecules in monolayers as a function of their intermolecular separation. He assumes that the total interaction between two molecules is given by the sum of the interactions of the functional groups of the molecules treated as though only homologous groups interacted. He derives expressions for the interaction between two homologous groups using the data compiled by Dunkel⁷ on the increment in the energy of cohesion of chain compounds caused by the addition of a given functional group to the chain. The reader is referred to ⁶ and ⁷ for details. The interaction energies between the groups forming long chain fatty acids are given by

$$\begin{aligned} v(\text{CH}_3) &= 5.61 \times 10^5 \rho^{-6} (\times 10^{-16} \text{ ergs}) \\ v(\text{CH}_2) &= 4.1 \times 10^5 \rho^{-6} \\ v'(\text{COOH}) &= 2.08 \times 10^4 \rho^{-3} \end{aligned} \quad (10)$$

where ρ is the distance between the two interacting groups. The total potential energy of interaction between a functional group in a given molecule and its homologues in all of the neighboring molecules in a monolayer is given, for the indicated groups, by

$$6.4 v(\text{CH}_3), \quad 6.4 v(\text{CH}_2), \quad \text{and} \quad 9.0 v'(\text{COOH}). \quad (11)$$

With these expressions and some knowledge or assumptions concerning the intermolecular orientations in monolayers, the quantity W and hence the energy barrier to evaporation can be estimated.

There are two different theories for the relative orientations and intermolecular separations of molecules in compressed fatty acid monolayers. The first postulates that the perpendicular distance between the vertical axes of neighboring molecules is constant during compression between the points A and B of FIGURE 3 and that the decrease in molecular area is caused by the tilting of the vertical axes away from the water surface as compression is applied. Presumably, the vertical axes are perpendicular to the water surface at B and make some angle with the surface at A . The second theory states that the decrease in area, as the film is compressed, is attributable to a decrease in the intermolecular separation. Dervichian⁸ has suggested that, at A , the molecules have a helicoidal configuration but at B are straight chains. Vold,⁹ on the other hand, has shown that the molecular areas at A and B may be attributable to the steric properties of the carboxyl end group, which in cross section is not a symmetrical group. At A , the molecular separation would just allow rotation about the perpendicular to the water surface while at B contact of the acid groups of neighboring molecules would prevent rotation. On either mecha-

nism, the transition between the states *A* and *B* is satisfactorily explained by the theory of monolayers of Joly.¹⁰

Although there is no evidence favoring either of these theories and excluding the other, the second is the more appealing if only because of the greater detail with which it has been expounded.

Nevertheless, assuming the validity of the proposed mechanism for penetration, only the "tilted chain" theory will accommodate the experimental results. Since specific resistance was found to be independent of the molecular area of the monolayer, the energy barrier *E* must be also, and since the energy barrier depends on the distance between molecules in the film, the intermolecular separation must be independent of the molecular area, which is a property only of the "tilted chain" theory.

Adopting this theory and assuming for purposes of simplifying the calculations that the chains are rigidly parallel, the energy barrier can be calculated from the relations 10 and 11 and is given by

$$E = W = 9.2 \times 10^5 (4.1n - 2.6)\rho^{-6} + 26.9 \times 10^4 \rho^{-3} \text{ cal./mole.}$$

Here *n* is the number of carbon atoms in the fatty acid molecule. The intermolecular separation ρ is 4.8 Å. if the area per molecule at *B* is taken as 20.0 Å². TABLE 2 shows the results of this calculation for the three acids under *E*_{calc}.

Actually, the Joly relations do not apply when the chains project at some angle to the water surface other than 90°, because the separation of homologous groups does decrease between the points *A* and *B* of the pressure-molecular area isotherm. Since the perpendicular distance between the chains does not change, however, if the interaction is considered to be between complete molecules, the interaction energy should be about the same for film molecules perpendicular to the surface and at some angle to the surface.

The experimental values for the specific resistances of the three acids are tabulated in column 2 of TABLE 2. The experimental energy barrier can be calculated by substituting in EQUATION 9. These values are listed in the column labeled *E*_{exp}. The increment corresponding to one —CH₂— group, calculated from the Joly relations, is compared in column 7 with the values determined from the experimental data tabulated in column 6. Columns 5 and 8 show the ratios of the calculated to the experimental energies.

The calculated values are consistently 10 per cent larger than the experimental values, both for the total energy barriers and the increment. A number of possible modifications of the theory could be made which would eliminate

TABLE 2
THEORETICAL AND EXPERIMENTAL ENERGY BARRIERS

Acid	<i>r</i>	<i>E</i> _{exp}	<i>E</i> _{calc}	<i>E</i> _c / <i>E</i> _e	Δ <i>E</i> _{exp}	Δ <i>E</i> _{calc}	Δ <i>E</i> _c /Δ <i>E</i> _e
C ₁₈	2.20	6979	7650	1.10			
C ₁₉	3.50	7265	7950	1.10	286	302	1.05
C ₂₀	5.45	7539	8250	1.10	274	302	1.10

(Energy values in calories/mole.)

this difference, but in view of the limited data and the uncertainty in the validity of the Joly calculations, which at best is probably not less than 10 per cent, a refinement of the theory is unwarranted. The agreement is sufficient to substantiate the principle of the theory of the resistance of monolayers to evaporation.

Review of Previous Work

The first attempt at determining the influence of surface films on the evaporation of water was that of Hedestrand¹¹ in 1924. Hedestrand tried to measure this influence by studying the effect of a film on the rate of evaporation of water into a current of streaming air. He concluded that monolayers were without effect on the evaporation rate. In view of subsequent work, this conclusion is, of course, untenable. The explanation is clear. The maximum rate at which water vapor was swept away by the air current was 3×10^{-4} g./min. per cm.² of clean water surface. This is four times smaller than the rate observed in the present work. It indicates that the resistances of Hedestrand's system, exclusive of the surface resistance, were extremely large, and, as a result, any changes noted in the resistance of the water surface caused by the film were obscured by the large total resistance. Furthermore, his films were necessarily contaminated with impurities since his experimental system was such that the water surface could not be swept clean before spreading a monolayer. Hedestrand used dilute benzene spreading solutions. It is evident from the present work that these factors of technique produce complications which almost eliminate the resistance to evaporation of any monolayer.

Rideal¹² was the first to show that monomolecular films have a marked influence on the evaporation rate of water. He used an inverted U tube to measure the effect. One arm contained the water at room temperature upon which the monolayer was formed. The other arm was cooled in an ice bath, and the system was evacuated. By comparing the rates of condensation in the cold arm with and without a surface film on the water, he could determine qualitatively the effect of the film on evaporation. By using a vacuum, he almost eliminated the resistance caused by the separation between the evaporating and condensing surfaces and consequently magnified the effects attributable to the resistance changes of the water surface. He found that stearic acid monolayers, for example, reduced the rate of condensation in the cold arm by about 28 per cent. His system is not tractable to analysis, so that no quantitative statement can be made about the effect of the film on the actual rate of evaporation, but he clearly showed that an effect of considerable magnitude does exist.

The next attempt was made by Sebba and Briscoe¹³ in 1940. It was their intention to establish the effect of surface pressure on the decrease in evaporation rate caused by monolayers. Therefore, they made their measurements with the water contained in the trough of a surface balance. They passed a stream of dry air inwards from the edges of a metal disk suspended above the water surface. The air, after passing over the surface, was withdrawn through an opening at the center of the disk and the amount of water vapor picked up was determined.

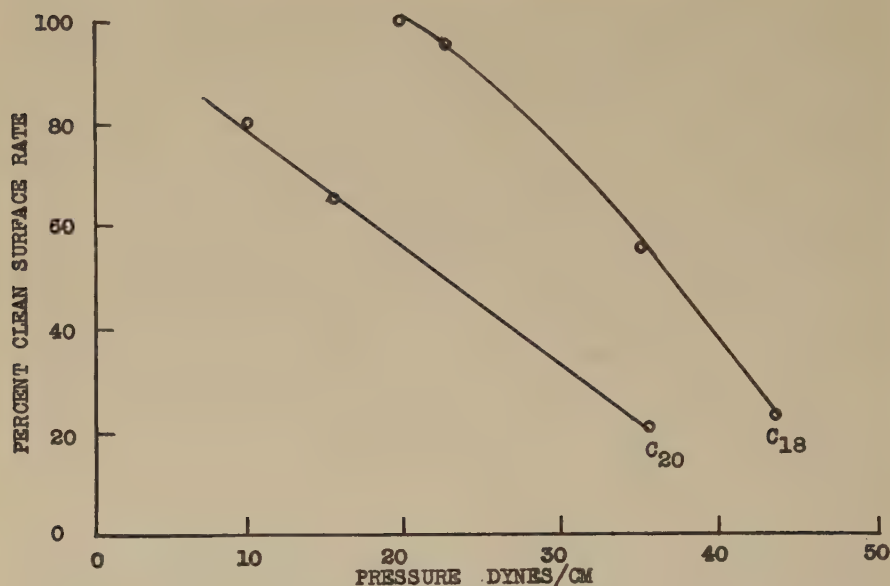


FIGURE 6. Effect of C_{18} and C_{20} acid monolayers on rate at which water vapor is swept away by streaming dry air, from data of Sebba and Briscoe.

Sebba and Briscoe plotted the fractional decrease in the rate at which water vapor is swept away, owing to the presence of the film, as a function of surface pressure. Curves 1 and 2 of FIGURE 6 show their results for the C_{18} and C_{20} acids. With the C_{18} acid, no decrease in the rate of uptake was found until the pressure reached about 25 dynes/cm. As the film was further compressed, the rate continuously decreased, and at 45 dynes/cm. was only about 30 per cent of the clean surface rate. With the C_{20} acid, the effect was greater and began with the first compression of the film. From these results, they concluded that the relative decrease in the rate of evaporation was strongly dependent on the surface pressure of the film.

Sebba and Briscoe suggest, as an explanation of this result, that the monolayers should be treated as solutions composed of water molecules and film molecules. At low surface pressures, the concentration of water molecules in the monolayer would be relatively large, and the evaporation would be little influenced by the film. With increasing pressure, water molecules would be expelled, and the rate of evaporation would decrease. Such a mechanism might give rise to the sort of curves they obtained. Since they used very dilute benzene spreading solutions, however, the characteristics of their curves are obviously attributable to the occlusion of large quantities of solvent molecules. With the C_{18} acid, for example, the fact that no effect on the evaporation rate was detectable until the pressure reached 25 dynes/cm. indicates that the concentration of occluded benzene was so large that it was only appreciably reduced by the very rapid rate of expulsion to be expected at high surface pressures.

Sebba and Briscoe also reported an aging effect. When they allowed the C_{20} monolayer to stand on the water surface for 20 hours at zero pressure and then compressed it to 45 dynes/cm., they found that the film had no effect on the evaporation rate, whereas in a normal determination the film at this pressure decreased the rate of water vapor uptake by 95 per cent. On the other hand, if the surface pressure was maintained above zero during the aging period, there was only a slight decrease in the evaporation effect when the monolayer was compressed to 45 dynes/cm. These investigators suggested that this result was caused by the formation of dimers in the uncompressed film by the combination of two C_{20} molecules in such a way that the resulting complex consisted of parallel chains with a carboxyl group at each end. Such a molecule would be expected to lie flat on the water surface and exhibit less resistance to evaporation.

It is difficult to imagine a mechanism whereby the described dimer could be formed, and Sebba and Briscoe suggest none. Furthermore, even if the dimer were formed, it would certainly not result in a film having the compressibility properties of the normal acid monolayer, and it could not be expected to reduce the surface tension of water to a surface pressure of 45 dynes/cm. under any compression because of the large separation of the polar end groups.

A more straightforward explanation can be made in terms of occluded impurities. When the film was allowed to stand uncompressed, it became saturated with impurities, but when the pressure was held above zero, both the rate of absorption and the saturation concentration of impurities were greatly reduced. Consequently, the resistance was much less affected under the latter conditions.

As in previous work, the data of Sebba and Briscoe cannot be related to the actual rate of evaporation through the monolayer because the physical properties of their system do not permit an analysis of the transport of water from the water surface to the stream of dry air.

The first significant advance was made by Langmuir and Schaefer¹ in 1943. They devised the technique for measuring evaporation resistance used in the present work, *i.e.*, suspension of a solid desiccant above the water surface. Since this technique permits a direct relation between the actual evaporation rate and the rate measured above the surface, it eliminates the principle difficulty of previous workers. Their equation relating the actual and measured evaporation rates is, in application, essentially the same as EQUATION 5 of this paper. Their development, however, is not based on a mathematical analysis of the transport system but depends on postulating the Ohm's law analogy and inserting the specific resistance of the surface in the denominator of the simple diffusion equation

$$M/At = (w - w_0)/(b/D).$$

The validity of their result depends on the validity of the Ohm's law analogy, which was not demonstrated. There is also ambiguity concerning the physical meaning of their equation since they imply that the quantities w and w_0 are the actual water vapor concentrations at the water and desiccant surfaces, which is not the case.

In calculating specific resistance from absorption rates, Langmuir and Schaefer use the equation

$$r = (1 - w_0/w) [(Atw/M)_{\text{film}} - (Atw/M)_{\text{no film}}], \quad (12)$$

considering the relative humidity term (w_0/w) to be constant and independent of the temperature of the system. They determined (w_0/w) for calcium chloride from the slope of the plot of the equation

$$DAtw/M = [1/(1 - w_0/w)]b$$

and found the relative humidity to be 0.21. This value corresponds to the equilibrium between the hydrates of calcium chloride containing 4 and 6 molecules of water. It is in disagreement with the more precisely determined value for w_0 made in this laboratory. The equilibrium water vapor concentration corresponds, for both calcium chloride and lithium chloride, to their saturated solutions. Furthermore, the approximation that the relative humidity factor in EQUATION 12 is constant and independent of temperature is not valid for this technique because w and w_0 are water vapor concentrations at different temperatures.

Langmuir and Schaefer reported that the rate of absorption by the desiccant for consecutive runs decreased as the amount of water absorbed increased, stating that the rate was 7 per cent smaller after a total of 1.5 g. of water had been absorbed. This decrease in rate is not caused by a change in the absorbing quality of the desiccant, as assumed by these investigators, but results from variations in the desiccant surface temperature. If an absorption run is made with the initial temperature of the desiccant corresponding to room temperature, a rate of absorption v will be observed, and the final desiccant temperature will be considerably above the initial value. If, then, a second run is made without first cooling the desiccant to the initial temperature of the first run, a different rate v' will be measured, since the average temperature of the desiccant surface for the second run will be higher than that for the first. Since $(w - w_0)$ decreases as the desiccant temperature increases, it is seen from EQUATION 5 that v will be larger than v' . We find that, if the initial temperature for consecutive runs is the same, no decrease in rate is observed.

Langmuir and Schaefer did not make the correction to M arising from water vapor in the air. Since this correction has a magnitude of the order of 0.010 g., a considerable error is inherent in their resistance values, especially for large r values, where M is relatively small. This factor, plus their use of an incorrect value for w_0 and failure to take into account the temperature of the desiccant, all lend considerable uncertainty to their data. In view of a more serious source of error inherent in their spreading technique, however, this uncertainty is unimportant.

FIGURE 7 shows their results for the C_{13} and C_{20} acids. Curves a and b are for the two acids on a dilute sulphuric acid solution, and curve c is for the C_{20} acid on pure water. As might be anticipated from the shape of these curves, Langmuir and Schaefer used dilute benzene solutions for spreading their films. They are similar to our curve I of FIGURE 4 for which a similar technique was

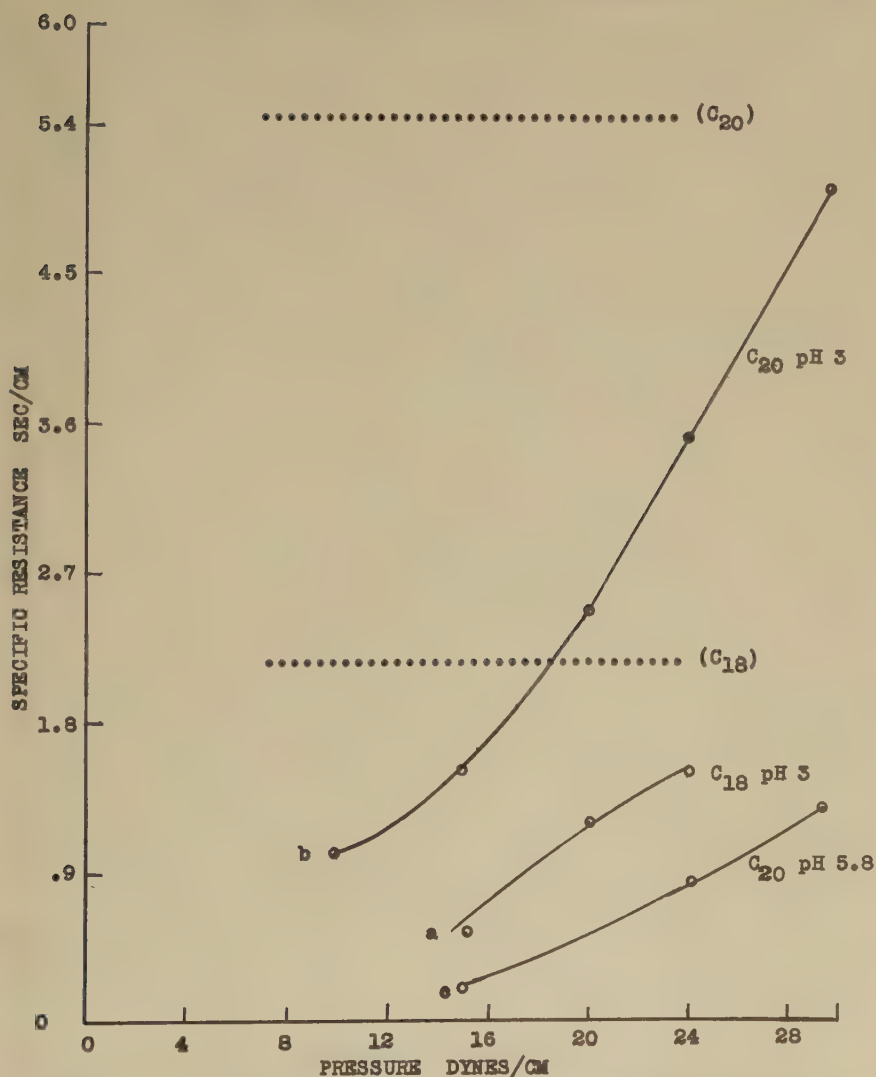


FIGURE 7. Data of Langmuir and Schaefer for monolayers of C_{18} and C_{20} acids on dilute sulfuric acid and of C_{20} acid on distilled water. Dotted curves show result of present work for these monolayers on distilled water.

used. Disregarding, momentarily, the pH of the substrate, the great difference between their results and those of the final technique of the present work, as indicated by the horizontal dotted lines for the C_{18} and C_{20} acids, clearly shows the effect attributable to solvation of benzene molecules in the films. In view of the contamination of their monolayers, no discussion of their numerous other curves will be given as they have little relation to the properties of pure monolayers.

Although, as yet, no systematic study of the effect of the pH of the water on the resistance has been made in this laboratory, the preliminary results obtained using dilute benzene and dilute petroleum ether spreading solutions indicate that there is no marked difference between the curves at pH 3 and at pH 5.8. The first measurements made on pure water at pH 5.8, before the solvation effect was observed, correspond to the pH 3 curves of Langmuir and Schaefer. In view of these facts, the great difference they report for the C_{20} acid on water at pH 3 and pH 5.8 would seem to be the result of some experimental aberration.

Langmuir and Schaefer presented their data as the plot of the logarithm of specific resistance as a function of pressure. They chose this representation because of their belief that the energy barrier was directly related to the film pressure. If, accordingly, α is written

$$\alpha = \exp(-E/kT),$$

the relation between specific resistance and surface pressure has the form

$$\log r = \text{constant} + \text{constant}'(p).$$

Some of the curves fit this relation fairly well, but this conformity is not a verification of a relation between E and p . It simply reflects the already noted exponential shape of the r - p plot for monolayers with occluded benzene molecules. The characteristics of such curves are determined by the rate of expulsion of occluded molecules with compression and by the change in specific resistance accompanying the change in the concentration of the occluded molecules. Finally, Langmuir and Schaefer were unable to account for the magnitude of E in terms of work against surface pressure as required by their relation between r and p .

The other reports on this problem add little to the historical development outlined here. There are several papers in the Russian literature of about 15 years ago and a recent publication by Sebba and Sutin.¹⁴ The latter use a cumbersome experimental arrangement to effect a partial vacuum around the surface balance and measure the rate of condensation on a "cold finger" above the water surface. Their results are similar to those of Sebba and Briscoe.

In conclusion, it must be emphasized that the data presented and the theoretical ideas proposed in this paper represent only an initial attack on the problem. There probably will be modifications in the theory of the resistance of monolayers as more comprehensive data are obtained. We intend to complete the work with an examination of the influence of the composition of the substrate on resistance and with an extension of the measurements to monolayers other than those of long chain fatty acids in the liquid condensed phase.

References

1. LANGMUIR, I. & V. J. SCHAEFER. 1943. *J. Franklin Inst.* **235**: 119.
2. BICHOWSKY, F. R. 1940. *Chem. & Met. Eng.* : 302.
3. HARKINS, W. D. 1952. *The Physical Chemistry of Surface Films*. Reinhold. New York, N. Y.
4. ALTY, T. 1931. *Proc. Roy. Soc. (A)* **131**: 554.

5. ANDERSSON, K. J. I., S. STALLBERG-STENHAGEN & E. STENHAGEN. 1944. The Svedberg (Mem. Vol.). Almqvist & Wiksells Boktryckeri Ab. Uppsala, Sweden.
6. JOLY, M. 1946. J. phys. radium. **7**: 83.
7. DUNKEL, M. 1928. Z. physik. Chem. (A) **138**: 42.
8. DERVICHIAN, D. G. 1949. Surface Chemistry. Butterworths Scientific Publications. London, England.
9. VOLD, J. 1952. J. Colloid Sci. **7**: 196.
10. JOLY, M. 1950. J. Colloid Sci. **5**: 49.
11. HEDESTRAND, G. 1924. J. Phys. Chem. **28**: 1244.
12. RIDEAL, E. K. 1925. J. Phys. Chem. **29**: 1585.
13. SEBBA, F. & H. V. A. BRISCOE. 1940. J. Chem. Soc. : 106.
14. SEBBA, F. & N. SUTIN. 1952. J. Chem. Soc. : 2513

CLASSICAL THEORY OF DIFFUSION AND THE OXIDATION OF METALS

By Earl A. Gulbransen

Westinghouse Research Laboratories, East Pittsburgh, Pa.

Introduction

The science of surface reactions has become a field of increasing importance in modern technology. Much work has been directed at understanding the mechanism of surface reactions, particularly in the field of oxidation of metals. This work has been profitable in the field of oxidation because of the rapid development of experimental methods for the study of the kinetics of the reaction and the structure of the oxide film.

One of the methods for testing the mechanism of a reaction is the comparison of predictions of the theoretical rate expression with experiment. This test was first made in oxidation of copper at high temperature by Wagner.¹ A theoretical expression was obtained based on the conductivity of the oxide and transference numbers of the ions and electrons involved in the reaction. The good agreement between theory and experiment showed the assumed model for the reaction to be the correct one. Unfortunately, the method has not been applied to other systems.

During the past ten years we have attempted an alternative approach² based on the diffusion picture of oxidation as stated by Mott together with the transition state theory of diffusion.⁴ This development gives an expression for the parabolic rate law constant in terms of absolute physical constants, two entropy of activation terms, and two heat of activation terms. An analyses of the experimental data⁵ showed a wide variation in the values for the sums of the heats, ΔH^* , and the sums of the entropies, ΔS^* , of activation. Thus, chromium, beryllium, and tungsten show $\Delta H^* > 40$ k.cal. per mole and ΔS^* values near zero or positive. However, most metals show ΔH^* values of 18 to 40 k.cal. per mole and ΔS^* of -10 to -30 entropy units.

Since the transition state theory of diffusion⁴ is based on partition functions, it is difficult to interpret the variable values obtained for the experimental entropy of activation term. In general, we have interpreted negative values of the entropies of activation in terms of diffusion at grain boundaries rather than through the crystal as a whole.⁶

Zener⁶ has recently restated the classical theory of diffusion in terms of the isothermal work necessary to move the diffusing atom or ion from its equilibrium position through the saddle configuration. This isothermal work can be expressed in terms of a heat of activation term and an entropy of activation term by familiar thermodynamic equations.

From simple thermodynamic considerations, the entropy of activation for diffusion can be shown to take only zero or positive values. The theory includes an expression for the calculation of this term from fundamental physical considerations. This fact has made it possible to calculate from theory the two entropy terms in the rate expression. This calculation, in turn, can be compared to the experimental value.

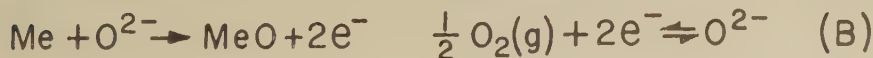
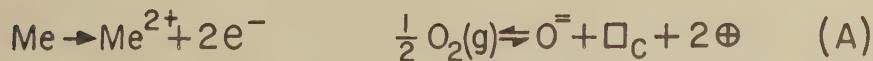
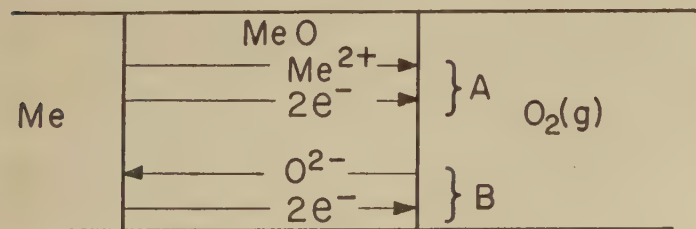
It is therefore possible to use this rate theory when combined with experimental rate of reaction data to determine the fundamental mechanism of the reaction.⁷ A review of the development of these ideas is the subject of this paper.

Oxidation Picture

The current picture of oxidation of metals assumes the growth of a uniform and homogenous oxide film which acts to limit further oxidation. The reaction occurs by one or more diffusion processes involving ions or atoms of the metal or of oxygen in the appropriate defects. It is also assumed that diffusion occurs in the lattice of the oxide and not at any limited zones in the crystal.

Using the simple picture of a homogeneous and uniform oxide film the parabolic rate law has been derived by Pilling and Bedworth,⁸ Tammann,⁹ Wagner,¹ Hoar and Price,¹⁰ and Mott.³ Experimental observations on the rate of oxidation of many metals in certain ranges of temperature and pressure and time show the parabolic rate law to be obeyed. This indicates that the simple picture is correct in regard to (1) the effect of average film thickness on the rate of reaction; and (2) the constant value for the number of interstitial ions or atoms, or of vacancies at the oxide-metal, or oxide gas interfaces over the thickness range tested.

FIGURE 1 shows a schematic picture of two alternative mechanisms for the oxidation of a metal. The first mechanism involves the diffusion of metal ions through vacancies in the oxide lattice, while the second involves the diffusion of oxygen atoms or ions through the oxide lattice. Before setting down an expression for the rate of oxidation, it is necessary to consider briefly the nature



$\oplus$ ELECTRON HOLE Me^{3+}

$\square_c$ Me^{2+} VACANCY

FIGURE 1. Oxidation scheme.

of defects in oxide lattices and the thermodynamics of the processes of formation.

Defects in Oxide Lattice

Koch and Wagner¹¹ have discussed the various types of defects which may occur in ionic lattices. These defects can be classed as stoichiometric and nonstoichiometric. The latter group are of interest to us in the general picture of oxidation. These defects occur by the solution of excess metal or oxygen in the oxide structure. Oxygen can dissolve in the oxide structure to form interstitial oxygen atoms or ions or in the form of vacant cation sites and electron holes at nearby lattice sites. In NiO the electron holes are assumed to be Ni^{3+} sites.

Metal can occur in excess by interstitial solution of metal atoms which may in turn be dissociated into metal ions and electrons or, by means of vacancies located at some of the normal oxide-ion sites, with the electrons located at positive ion sites.

FIGURE 2 shows the oxygen excess type in NiO. The NiO lattice is face-centered cubic. When one half mole of oxygen goes into solution according to the equation



one cation vacancy $\square_c$ is formed together with two positive holes $\oplus$ and one oxygen ion O^{2-} . If we assume that no electrons are added or taken away, the number of cation vacancies per cm.³, $n_{\square_c}$, equals one half the number of positive holes, $n_{\oplus}$.

NiO Lattice	Defect Equation
$\text{Ni}^{2+} \quad \text{O}^{2-} \quad \text{Ni}^{2+} \quad \text{O}^{2-} \quad \text{Ni}^{2+}$	$\frac{1}{2} \text{O}_2 \rightleftharpoons \text{O}^{2-} + \square_c + 2 \oplus$
$\text{O}^{2-} \quad \text{Ni}^{2+} \quad \text{O}^{2-} \quad \square_c \quad \text{O}^{2-}$	Equilibrium Constant
$\text{Ni}^{2+} \quad \text{O}^{2-} \quad \text{Ni}^{3+} \quad \text{O}^{2-} \quad \text{Ni}^{2+}$	$K = \frac{n_{\square_c} n_{\oplus}^2}{P_{\text{O}_2}^{1/2} N^3}$
$\text{O}^{2-} \quad \text{Ni}^{2+} \quad \text{O}^{2-} \quad \text{Ni}^{2+} \quad \text{O}^{2-}$	Conductivity
$\text{Ni}^{3+} \quad \text{O}^{2-} \quad \text{Ni}^{2+} \quad \text{O}^{2-} \quad \text{Ni}^{2+}$	$n_{\oplus} = K^{1/3} P_{\text{O}_2}^{1/6} N$
	Reaction Rate
	$n_{\square_c} = K^{1/3} P_{\text{O}_2}^{1/2} N$
$\square_c$ Cation Vacancy Ni^{3+} Positive Hole = $\oplus$	

FIGURE 2. Lattice defects NiO (vacancy type).

The equilibrium constant for the defect equation is

$$k = \frac{n_{\square_e}}{N} \cdot \left(\frac{n_{\oplus}}{N} \right)^2 \frac{1}{p_{\text{O}_2}^{1/2}} \quad (2)$$

Here N is the total number of sites per cm.³

$\frac{n_{\square_e}}{N}$ is the concentration of cation vacancies,

$\frac{n_{\oplus}}{N}$ is the concentration of positive holes, and

p_{O_2} is the pressure of oxygen in atmospheres.

We can now solve for the number of cation vacancies, since

$$n_{\oplus} = 2n_{\square_e}, \quad (3)$$

then

$$n_{\square_e} = \frac{N}{4^{1/3}} (p_{\text{O}_2})^{1/6} K^{1/3} \quad (4)$$

The number of positive holes is given by

$$n_{\oplus} = 2^{1/3} N (p_{\text{O}_2})^{1/6} K^{1/3} \quad (5)$$

The equilibrium constant can be given in terms of the standard free energy of formation by the equation

$$\Delta F^0 = -RT \ln K \quad \text{or } K = e^{-\Delta F^0/RT} \quad (6)$$

Substituting, we have

$$n_{\square_e} = \frac{N}{4^{1/3}} (p_{\text{O}_2})^{1/6} e^{-\Delta F^0/3RT} \quad (7)$$

and

$$n_{\oplus} = 2^{1/3} N (p_{\text{O}_2})^{1/6} e^{-\Delta F^0/3RT} \quad (8)$$

Since

$$\Delta F^0 = \Delta H^0 - T\Delta S^0 \quad (9)$$

$$n_{\square_e} = \frac{N}{4^{1/3}} (p_{\text{O}_2})^{1/6} e^{-\Delta H^0/3RT} e^{\Delta S^0/3R} \quad (10)$$

and

$$n_{\oplus} = 2^{1/3} N (p_{\text{O}_2})^{1/6} e^{-\Delta H^0/3RT} e^{\Delta S^0/3R} \quad (11)$$

These equations give expressions for the number of vacancies which are

useful in the diffusion process or for the number of electron holes responsible for conductivity.

FIGURE 3 shows the metal excess type in ZnO. The zinc atoms occupy interstitial positions and are dissociated into ions and electrons. The equations for forming singly- and doubly-charged zinc ions are given together with expressions for the number of electrons per cm.³ and the number of vacancies.

$$n_{e^-} = N(p_{O_2})^{-1/4} K^{1/2}, \quad (12)$$

$$n_{Zn^{+}} = N(p_{O_2})^{-1/4} K^{1/2} \quad (13)$$

Here n_{e^-} is the number of electrons formed per cm.³, $n_{Zn^{+}}$ is the number of zinc ions formed per cm.³, and N is the total number of sites per cm.³

For the reaction to form doubly-charged zinc ions we have the following equations:

$$n_{e^-} = 2^{1/3} N(p_{O_2})^{-1/6} K^{1/3} \quad (14)$$

$$n_{Zn^{++}} = \frac{N}{4^{1/3}} (p_{O_2})^{-1/6} K^{1/3} \quad (15)$$

Here $n_{Zn^{++}}$ is the number of zinc ions formed per cm.³

Since K is the equilibrium constant and it is related to the free energy of formation of the interstitial ion, K may be expressed in terms of ΔF^0 . Thus, for the formation of doubly charged ions, we have

$$n_{e^-} = 2^{1/3} N(p_{O_2})^{-1/6} e^{-\Delta F^0/3RT} \quad \text{or} \quad 2^{1/3} N(p_{O_2})^{-1/6} e^{-\Delta H^0/3RT} e^{\Delta S^0/3R} \quad (16)$$

$$n_{Zn^{++}} = \frac{N}{4^{1/3}} (p_{O_2})^{-1/6} e^{-\Delta F^0/3RT} \quad \text{or} \quad \frac{N}{4^{1/3}} (p_{O_2})^{-1/6} e^{-\Delta H^0/3RT} e^{\Delta S^0/3R} \quad (17)$$

These are fundamental expressions which apply to the formation of interstitial ions.

The simple picture of oxidation suggests that the quantities ΔH^0 and ΔS^0 are uniform throughout the lattice. This is not expected to be always true since it is recognized that stresses may occur at grain boundaries and at other regions in the metal crystal. For this reason, ΔH^0 and ΔS^0 may differ appreciably from the normal values with the rate of reaction at these areas being greatly favored. These local areas of stress have been shown to exist by Gulbransen, McMillan, and Andrew¹² in the growth of oxide crystallites at low pressures of oxygen and for high temperature.

Parabolic Rate Law

We now derive the parabolic rate law following the method and notation of Mott³ for the case when the formation and diffusion of electrons is easy compared to the formation and diffusion of ions or vacancies and when no special

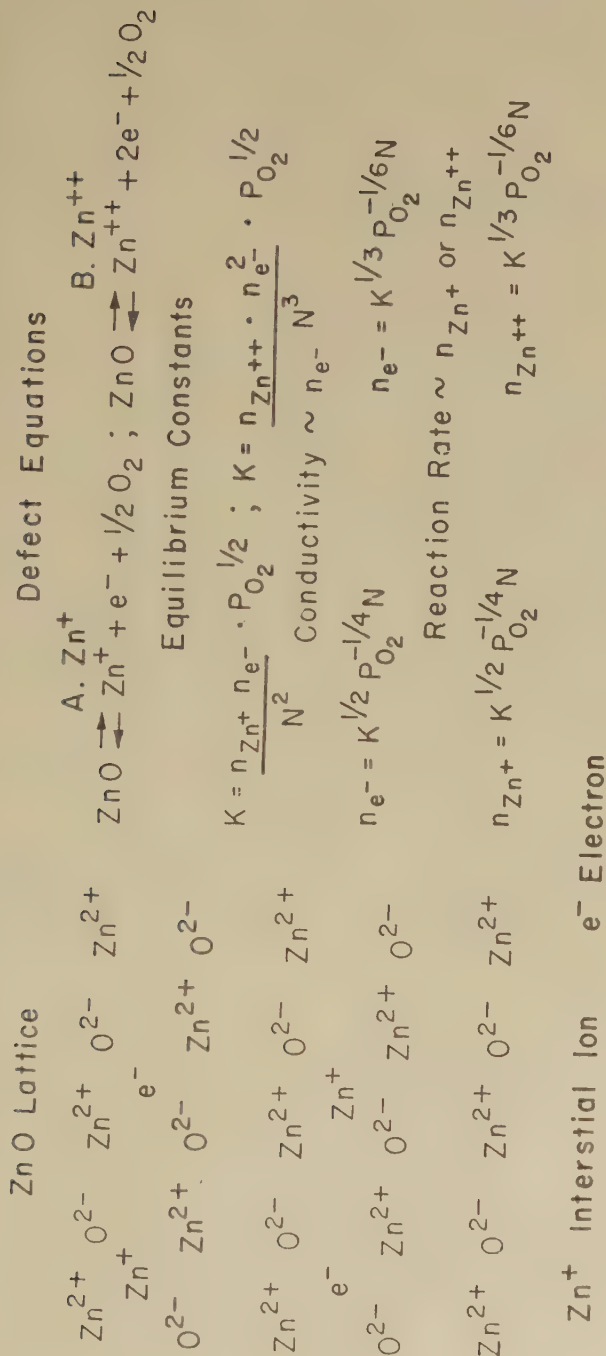


FIGURE 3. Lattice defects ZnO (interstitial type).

electrical field exists across the oxide film. Under these conditions the ion current j is given by

$$-j = -D \frac{dn}{dx} \quad (18)$$

Here, D is the diffusion coefficient and dn/dx is the concentration gradient. Integrating across the layer of thickness X_0

$$jX_0 = 2D(n_1 - n_2) \quad (19)$$

If Ω is the volume of the oxide per metal ion

$$\frac{dx_0}{dt} = j\Omega = 2D \frac{(n_1 - n_2)}{X_0} \Omega \quad (20)$$

or

$$X^2 = At, \quad A = 2\Omega D(n_1 - n_2) \quad (21)$$

Here, n_1 and n_2 are the number of vacancies or interstitial ions per cm.³ at the oxide gas and oxide metal interfaces.

To carry the theory further, we must consider the nature of the reaction for a particular metal-oxide system. Let us consider the special case of the oxidation of pure nickel. Nickel oxide is a p type of semiconductor with the vacancies formed by EQUATION 1. Since the rate of oxidation is proportional to the number of vacancies $n_{\square}$, we can substitute equation 10 in equation 21. Since $n_2 = 0$ at the metal oxide interface, the parabolic rate law constant becomes

$$A = \frac{2\Omega D}{4^{1/3}} N(p_{O_2})^{1/6} e^{-\Delta F^0/3RT} \quad (22)$$

Classical Theory of Diffusion

The classical theory of diffusion as restated by Zener⁶ is based on the random walk problem, and the theory gives a relationship between the diffusion coefficient D , the vibration frequency ν of the lattice in the jump direction and the isothermal work ΔF^* for diffusion,

$$D = \gamma a^2 \nu e^{-\Delta F^*/RT} \quad (23)$$

Here γ is a constant characterizing the nature of the atom jumps and a is the distance between jumps.

Substituting this value of D in equation 22 we have

$$A = 2 \frac{\gamma}{4^{1/3}} a^2 \Omega \nu N(p_{O_2})^{1/6} e^{-\Delta F^0/3RT} e^{-\Delta F^*/RT} \quad (24)$$

Since

$$\Delta F^0 = \Delta H^0 - T\Delta S^0 \quad \text{and} \quad \Delta F^* = \Delta H^* - T\Delta S^* \quad (25)$$

$$A = 2 \frac{\gamma a^2 \Omega}{4^{1/3}} \nu N(p_{O_2})^{1/6} e^{(\Delta S^0/3 + \Delta S^*)/R} e^{-(\Delta H^0/3 + \Delta H^*)/RT} \quad (26)$$

Here ΔS^0 = standard entropy of formation of vacancies

ΔH^0 = standard heat of formation of vacancies

ΔS^* = entropy of activation of diffusion

ΔH^* = heat of activation of diffusion

R = gas constant

T = absolute temperature

Absolute Rate of Oxidation

The theoretical expression for the parabolic rate law constant A is used to establish a mechanism for the reaction in the following way. A mechanism of formation of the defects by which diffusion occurs is assumed. This determines the particular expression for the parabolic rate law constant A . The entropy of formation of the defects is calculated, and the other quantities in the rate expression are determined from the structure of the particular oxide existing on the metal, with the exception of the heat of formation of defects and the heat of activation of diffusion. The sum of the latter terms are determined experimentally from the temperature coefficient of the parabolic rate law constant. The theoretical rate expression for A at a temperature where the parabolic rate law applies is now compared to the experimental value. Agreement indicates the assumed picture is the correct one.

Calculation of Absolute Rate of Oxidation of Nickel

The following steps are needed to evaluate the parabolic rate law constant A .

θ_D = Debye temperature

$$(1) \quad \nu = \frac{\theta_D k}{h} \quad k = \text{Boltzmann's constant}$$

h = Planck's constant

(2) γ , a , ν and N are determined from the structure of NiO.

(3) $\Delta H^0 + \Delta H^*$ are determined as a sum from the experimental plot of $\log A$ vs. $1/T$. FIGURE 4 shows a plot of $\log A$ vs. $1/T$ for NiO. For NiO ($\Delta H^0 + \Delta H^*$) is equal to 41,200 calories per mole.

(4) ΔS^0 is calculated from classical theory of ionic crystals and thermodynamic data. The change in entropy for the condensation of oxygen into the lattice can be estimated from two considerations:

(a) Entropy change assuming no distortion of the nickel oxide lattice.

(b) Entropy change of distortion of NiO lattice owing to the formation of the vacancy and the two positive holes.

This calculation has been made in detail for NiO by Gulbransen and Andrew.⁷

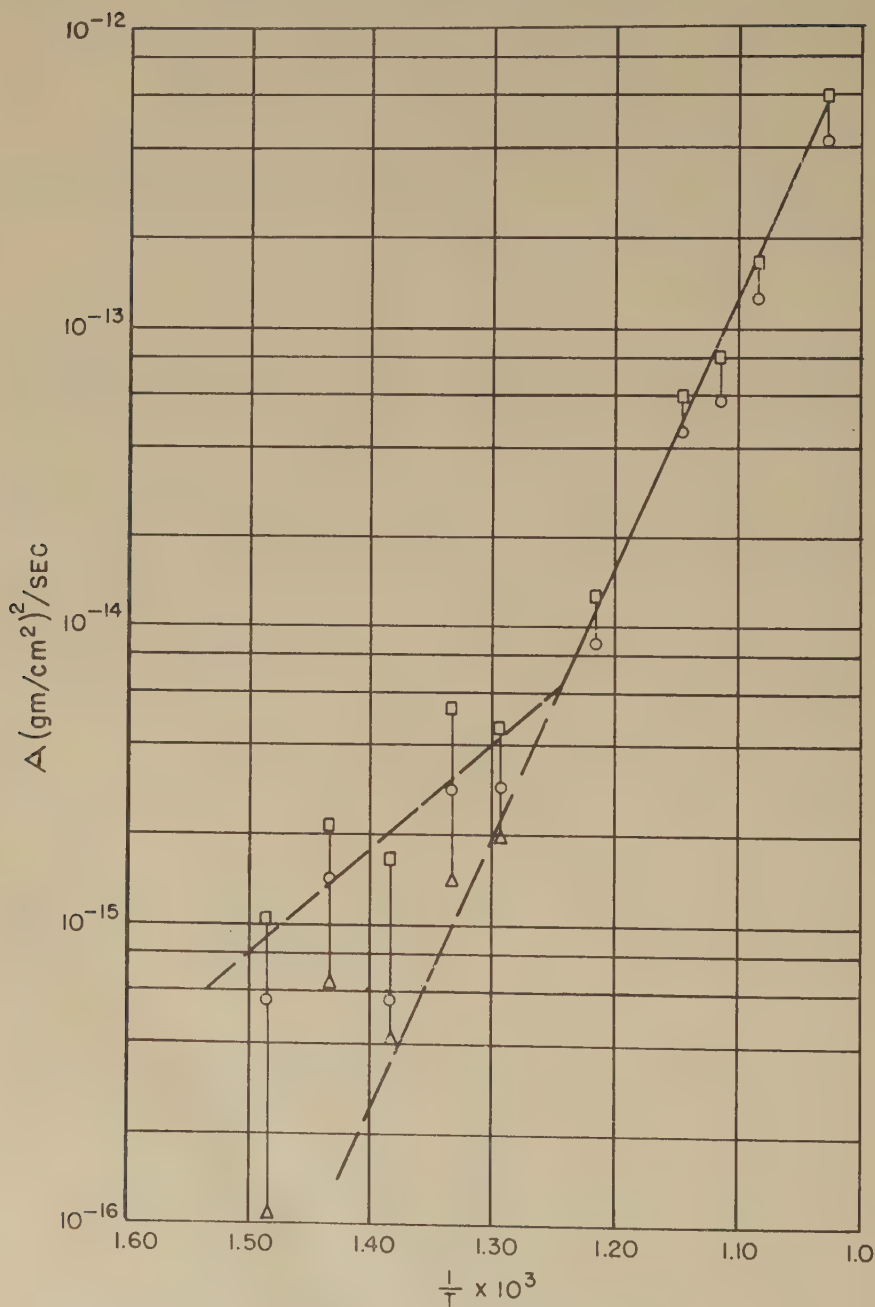


FIGURE 4. Oxidation of semidegassed nickel. Log plot A versus $\frac{1}{T} \times 10^3$. $\square$ 1 to 2 hrs. $\circ$ 3 to 6 hrs. $\triangle$ 24 to 30 hrs. $H^* = 41,200$ Cals. Mole.

At 1000° K for NiO

$$S_{1/2O_2(g)} = 29.10 \text{ calories per degree}$$

$$S \text{ of } O^= \text{ in NiO} = 9.04 \text{ calories per degree}$$

$$\Delta S \text{ of distortion} = 2.4 \text{ calories per degree}$$

Therefore $\Delta S^0 = -22.5$ calories per degree. Since three defects are formed in the process

$$\Delta S^0 = -7.5 \text{ calories per degree per mole of defects.}$$

Calculation of ΔS^*

The entropy of activation of diffusion has a positive value. This can be seen from a consideration of the equation

$$\Delta S^* = -\frac{\partial \Delta F^*}{\partial T}$$

ΔF^* is the isothermal work. $\partial \Delta F^* / \partial T$ would be expected to have a negative value, therefore

$$\Delta S^* = +$$

Theoretical values for ΔS^* can be evaluated from an equation by Zener⁶

$$\Delta S^* = \lambda \frac{\beta H^*}{T_m}$$

λ represents the fraction of the free energy of activation or work that goes into straining the lattice.

ΔH^* is the heat of activation of diffusion.

T_m is the melting temperature of NiO.

$$\beta = -d(\mu/\mu_0)/d(T/T_m).$$

μ is the elastic modulus, μ_0 is the extrapolated value of μ at $T = 0$.

β can be estimated from the statistical thermodynamic expression for C_p

$$C_p = \frac{\partial}{\partial T} \left\{ RT^2 \left(\frac{\partial \ln Q_v}{\partial T} \right)_v \right\} \quad Q_v = \frac{kT}{h}$$

Differentiating

$$C_p = R \frac{\partial}{\partial T} \left\{ T - T^2 \frac{d \ln \nu}{dT} \right\}$$

According to Einstein, a simple relation exists between the compressibility χ of a solid and the characteristic frequency ν_E

$$\nu_E \sim \frac{1}{\chi}^{1/2} \quad \text{Since } \chi \sim 1/\mu$$

$$\nu_E \sim \mu^{1/2}$$

Therefore

$$C_p = \frac{R\partial}{\partial T} \left\{ T - \frac{T^2}{2} \frac{d \ln \mu}{dT} \right\}$$

Rearranging and substituting

$$C_p = R + \frac{RT}{T_m} \frac{\mu_0}{\mu} \beta$$

For three degrees of freedom

$$C_p = 3R + 3R \frac{T}{T_m} \frac{\mu_0}{\mu} \beta$$

Since $\mu = \mu_0$ at $T = 0$ and $\mu = 0$ at T_m

$$\frac{\mu}{\mu_0} = 0.6 \text{ at } 1000^\circ K$$

Therefore $\beta = 0.42$. To calculate ΔS^* we need to assume what fraction ΔH^* is of the sum $(\Delta H^* + \Delta H^0)$. If we assume ΔH^* is $\frac{1}{2}(\Delta H^* + \Delta H^0)$ then for NiO

$$\Delta H^* = 20,600$$

and

$$\Delta S^* = 2$$

Calculation of Absolute Rate of Oxidation of Cobalt

Although the same methods are used, the calculations are more difficult for cobalt since the oxide film consists of two layers, Co_3O_4 and CoO , and good thermodynamic information is not available. The present calculations will test the assumptions of a vacancy diffusion mechanism involving the CoO layer.

The vacancies are assumed to form according to EQUATION (1) with the oxygen being furnished by the equation



The factor $p_{\text{O}_2}^{1/6}$ in EQUATION (4) can be expressed from thermodynamics as

$$p_{\text{O}_2}^{1/6} = e^{\Delta S_1^0/3R} e^{-\Delta H_1^0/3RT} \quad (28)$$

Here ΔS_1^0 and ΔH_1^0 are the standard entropy and heat of reaction of EQUATION (27). Both can be calculated from an equation given by Watanabe.^{13,14} Using the relation

$$\Delta S_1^0 = -\frac{\partial \Delta F^0}{\partial T}, \quad \Delta S_1^0 = 29.8 \text{ calories per mole per degree}$$

The experimental entropy and heat of activation for the oxidation of cobalt contains three factors.

TABLE 1
 SUMMARY OXIDATION DATA

	Nickel (700°C.)	Cobalt (400°C.)
Oxide.....	NiO	CoO + S.A. Co ₃ O ₄ *
Assumed type of defects.....	Vacancies	Vacancies
$\left(\frac{\Delta H_1^0}{3} + \frac{\Delta H^0}{3} + \Delta H^*\right)$ cals./mole (exptl.).....	41,200 $\left(\text{Log } K \text{ vs. } \frac{1}{T}\right)$	25,000 $\left(\text{Log } K \text{ vs. } \frac{1}{T}\right)$
$\left(\frac{\Delta S_1^0}{3} + \frac{\Delta S^0}{3} + \Delta S^*\right)$ cals./mole/ deg. (exptl.).....	-6.0	-10.3
$\frac{\Delta S_1^0}{3}$ (theory).....	0	9.9
$\frac{\Delta S^0}{3}$ (theory).....	-7.5	-7.7
ΔS^* (theory).....	2	1
ΔS^* (exptl.).....	1.5	-12.5
Parabolic constant (exptl.).....	2.04×10^{-13} cm. ² /sec.	6.21×10^{-13}
Parabolic constant (theory).....	2.70×10^{-13} cm. ² /sec.	1.8×10^{-10}

* Small amount.

Summary of Oxidation Data

TABLE 1 summarizes the application of the absolute rate theory to the oxidation of nickel and cobalt. For nickel, the agreement between the experimentally determined ΔS^* , the entropy of activation of diffusion and the calculated value is excellent considering the errors involved. The experimental and theoretical parabolic rate constants are also compared.

For cobalt, the agreement is poor. We may conclude that the assumed mechanism for the reaction is not the correct one and that other mechanisms having negative entropy terms should be tested. Unfortunately, in the case of cobalt, the necessary experimental information on the oxide layer thicknesses is lacking and calculations at this time would be speculative.

 TABLE 2
 GENERAL CONSIDERATIONS OF DEFECT FORMATION

A. Negative entropy change

1. Vacancy formation $\frac{1}{2} \text{O}_2 \rightleftharpoons \text{O}^{2-} + \square_e + 2 \oplus, \Delta S^0 = -6 \text{ to } -10$
2. Interstitial oxygen $\frac{1}{2} \text{O}_2 \rightleftharpoons \text{O} \text{ (interstitial)}, \Delta S^0 = -20 \text{ to } -25$
3. Adsorption of O_2 , $\text{Ni (inter.)} + \frac{1}{2} \text{O}_2 \rightleftharpoons \text{O}^- \text{ (Ads.)} + \text{Ni}^{++}, \Delta S^0 = -20 \text{ to } -25$

B. Near zero entropy change

1. Interstitial metal atom, $\text{Ni (metal)} \rightleftharpoons \text{Ni (interstitial)}, \Delta S^0 \approx 0$
2. Interstitial metal ion, $\text{Ni (metal)} \rightleftharpoons \text{Ni}^{++} \text{ (interstitial)}, \Delta S^0 \approx 0$
3. Vacancy formation $\text{M}_3\text{O}_4 + \text{M}^{2+} \rightleftharpoons 4\text{MO} + \square_e + 2 \oplus, \Delta S^0 \approx 0$

It is of interest to show what range of values the entropy term for different processes of defect formation may take. These values are summarized in TABLE 2. Several cases are known in the literature where experimental entropies of activation of -20 to -25 entropy units occur. These cases are usually associated with reactions having low heats of activation. This suggests that the formation of defects with a large negative entropy term involves a process with a small heat of formation. The converse of this is also true. Processes with zero or positive entropy of formation of defects are processes with high heats of formation of defects.

TABLE 2 can be used to study the mechanism of reaction in the following way. Calculations are made for the several alternative processes of forming defects. The particular process which when combined with the entropy of activation of diffusion gives the correct experimental entropy term is judged to be the correct mechanism for reaction.

Acknowledgment

The author wishes to express his appreciation to Doctor C. Zener and Doctor C. Wagner for their helpful discussions and suggestions.

References

1. WAGNER, C. 1933. *Z. physik. Chem. B.* **21**: 25.
2. GULBRANSEN, E. A. 1943. *Trans. Electrochem. Soc.* **83**: 301-17.
3. MOTT, N. F. 1940. *Trans. Faraday Soc.* **36**: 472-483.
4. GLASSTONE, LAIDLER & EYRING. 1941. *The Theory of Rate Processes*. McGraw-Hill, New York, N. Y.
5. GULBRANSEN, E. A. 1954. *Proceedings of the Gothenburg Conference on Solid State Reactions, 1952*. Gothenburg, Sweden.
6. ZENER, C. 1951. *J. Applied Phys.* **22**(4): 372-375.
7. GULBRANSEN, E. A. & K. F. ANDREW. 1954. *J. Electrochem. Soc.* **101**: 128.
8. PILLING, N. B. & R. E. BEDWORTH. 1923. *J. Inst. Metals.* **29**: 529.
9. TAMMANN, G. 1920. *Z. anorg. u. allgem. Chem.* **111**: 78.
10. HOAR, T. P. & L. E. PRICE. 1938. *Trans. Faraday Soc.* **34**: 867.
11. KOCH, E. & C. WAGNER. 1938. *Z. physik. Chem. B.* **38**: 295.
12. GULBRANSEN, E. A., W. R. McMILLAN & K. F. ANDREW. *J. Metals*. In press.
13. WATANABE, M. 1934. *Science Reports Tôhoku Imp. Univ. First Ser.* **23**: 89.
14. GULBRANSEN, E. A. & K. F. ANDREW. 1951. *J. Electrochem. Soc.* **98**: 241-51.

PASSIVITY AND ADSORPTION

By Herbert H. Uhlig

*Corrosion Laboratory, Department of Metallurgy, Massachusetts Institute of Technology,
Cambridge, Mass.*

Forces at the surface of a metal are merely extensions of those forces within the crystal serving to hold it together. At the same time, the average number of neighbors of surface atoms are only half those of atoms underneath, leading to unbalance of forces at the surface. This situation accounts for a certain marked affinity of the surface for its environment, and also for a stronger bonding of surface atoms to their underlying neighbors and to themselves than is common to the lattice as a whole. It is well known, for example, that the strength of bonds between atoms increases as the number of neighbors decreases.^{1, 2} Pauling³ has also shown that, as the strength of bonds increases, atoms move closer to each other, or the interatomic distance decreases. This tendency immediately leads to the deduction that the distance between surface atoms with stronger-than-usual bonding to each other is less than the interatomic spacing of the lattice. If so, there must be a certain tendency of the surface atoms to compress atoms underneath, giving rise to the familiar surface tension easily demonstrated in the case of liquids, but not so easy to demonstrate in the case of solids. A further deduction can be made that if components of the environment adsorb on the surface of a metal, this adsorption supplies in effect additional neighbors, thereby decreasing the mutual bonding between surface atoms and, hence, reducing surface tension. This effect is a natural consequence of the environment satisfying some of the affinity of the metal, physical or chemical, for its environment.

Several experiments confirm these deductions. First, Hofmann and Wilm⁴ showed that small graphite particles have a smaller lattice parameter than large graphite particles, the reduction being caused, presumably, by an overall greater compression of the solid particles when the ratio of surface to volume is large. Second, Benedicks⁵ showed that a platinum wire 0.05 mm. in diameter and 246 mm. long expanded 0.0008 per cent on exposing to water, returning to its original dimensions on drying. The adsorbed water molecules reduced surface tension of solid platinum to an extent that was measurable in a fine wire. Similar results were obtained with silica fibres using solvents such as ether, xylol, and alcohol. Third, Buttner, Funk, and Udin⁶ measured the surface tension of weighted silver wires at elevated temperatures in both helium, where the value was 1140 dynes/cm. at 900° C, and in air, where the value was only 450 dynes/cm. The decrease of surface tension in air is attributed to the adsorption of oxygen on the silver, directly confirming that such a reduction takes place through satisfaction of surface affinities.

Properties of Adsorbates on Metals

Langmuir⁷ showed that surface forces of solids are relatively short range and are effectively cut off by as little as one atomic or molecular layer of adsorbate

on the surface. When strong affinities exist between the solid and the environment, adsorption, if it occurs at all, is called chemisorption, indicative of strong bonding. Subsequent layers of adsorbate on chemisorbed layers are much less tightly held because surface forces are so effectively diminished by the first layer and, hence, the bonding tends towards the weaker so-called physical adsorption.

In view of the fact that physical adsorption is rapid, whereas chemisorption may be slow, it is sometimes found that a species will first physically adsorb, followed by a transition of the first layers of adsorbate to a chemically bonded or chemisorbed layer. This tendency is illustrated by the diagram of FIGURE 1 where physically adsorbed molecular hydrogen acquiring the activation energy E becomes chemisorbed as hydrogen atoms, thereby falling into a lower energy trough and being located on the average a closer distance from the metal surface. On heating such a surface, both atomic hydrogen is emitted following the upper energy curve, as well as molecular hydrogen following the lower energy curve.^{8, 9}

From what has been said above, it is obvious that chemisorbed species more completely satisfy surface forces than do physically adsorbed species and, hence, the former are expected to have greater effect on surface tension and related phenomena. It is perhaps also obvious that a chemisorbed monolayer, by better satisfying surface affinities, renders such a surface less attractive to its environment than a physically adsorbed layer and more appreciably affects observed chemical properties of the surface. This relation accounts for the phenomena associated with passivity of metals typified by the reduced reac-

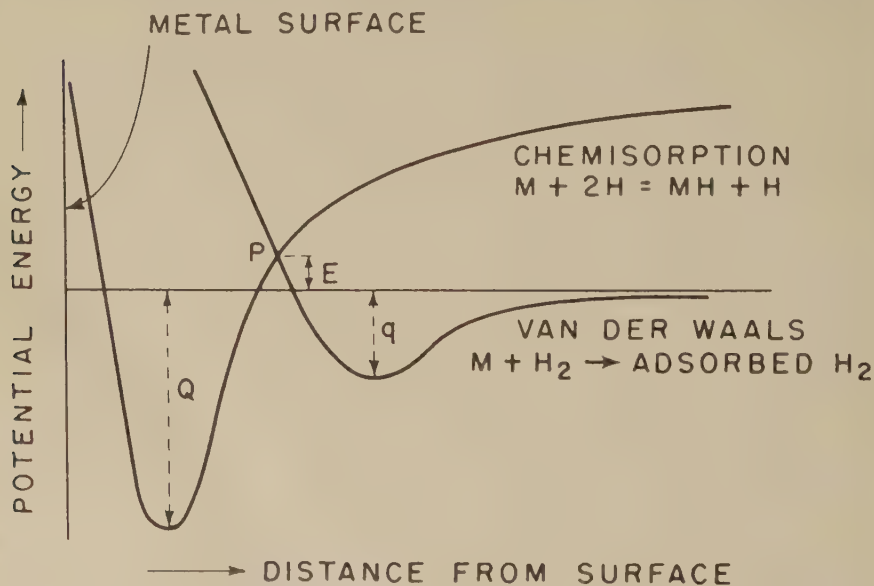


FIGURE 1. Potential energy diagram for physical adsorption and chemisorption.

tivity of iron immersed in concentrated nitric acid, or in chromate solutions, and by the various stainless steels.

Langmuir,^{7, 10} who first suggested that passivity derives from satisfaction of surface forces, showed that oxygen chemisorbed on tungsten is much less reactive with hydrogen than either molecular oxygen or oxygen of the WO_3 lattice. He showed, furthermore, that chemisorbed oxygen atoms become negatively charged, thereby drastically reducing the rate of electron emission from heated tungsten wires. Since the work function of a metal is reduced in accord with the equation:

$$\Delta\phi = 4\pi n\mu \quad (1)$$

where $\Delta\phi$ is the change of work function caused by n adsorbed oxygen atoms, and where μ is the dipole moment equal to the center of gravity of negative charge in the adsorbate atom multiplied by twice the distance of the center of gravity to the metal surface.

A negative charge of this kind on the surface of the metal and accompanying reduced reactivity of surface atoms also account for a more noble potential of the metal as measured in aqueous media. The change of potential is proportional to the number of adsorbate species and it accordingly follows an equation similar to EQUATION (1) above, the metal becoming increasingly noble and finally reaching a maximum value when the monolayer is approximately complete.

The amount of gas adsorbing on a surface as a function of pressure was shown by Langmuir¹¹ to be equivalent to:

$$x = \frac{ax_m p}{1 + ap}$$

where x is the amount of gas adsorbed per unit area of surface at pressure p , x_m is maximum adsorbate corresponding to high values of p , and a is a constant related to heat of adsorption. Using Henry's Law and substituting the concentration of solute C for the pressure p , and also making use of EQUATION 1, above, it follows that:

$$\frac{C}{\Delta E} = \frac{C}{\Delta E_m} + \frac{1}{a'\Delta E_m}$$

where ΔE_m is the maximum potential change and now a' is related to heat of adsorption. Hence, plotting $C/\Delta E$ versus C produces a straight line when conditions of the Langmuir expression are obeyed. This tendency is essentially the case when iron is exposed to increasing concentrations of chromate as shown in FIGURES 2 and 3, or when stainless steels are exposed to increasing pressures of oxygen as shown in FIGURES 4 and 5.¹² Optimum inhibition in the case of chromate occurs, as shown by Robertson,¹³ at a minimum of 10^{-3} molar, which is the approximate concentration for maximum potential change or monolayer formation of chromate ions. The fact that the Langmuir adsorption isotherm is obeyed in these cases adds weight to the proposal that passivity in metals is related to the satisfaction of surface affinities by monolayers of chemisorbed species.

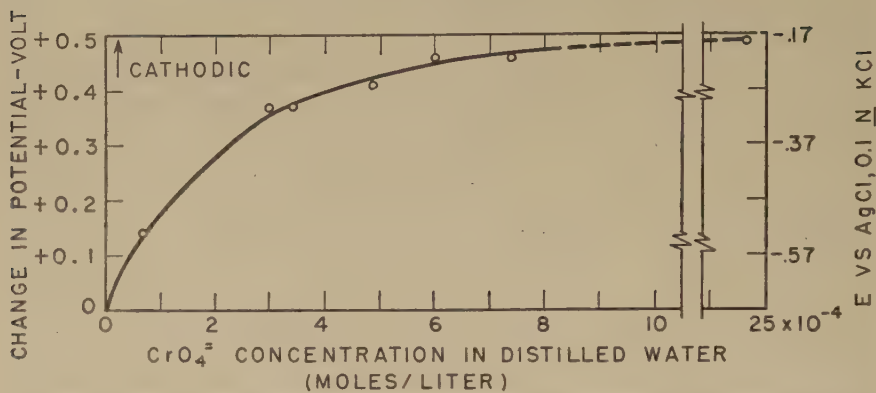


FIGURE 2. Effect of $\text{CrO}_4^{=}$ additions on the electrode potential of electrolytic iron in distilled water. $T = 25.0 \pm 0.2^\circ\text{C}$. Aerated solutions.

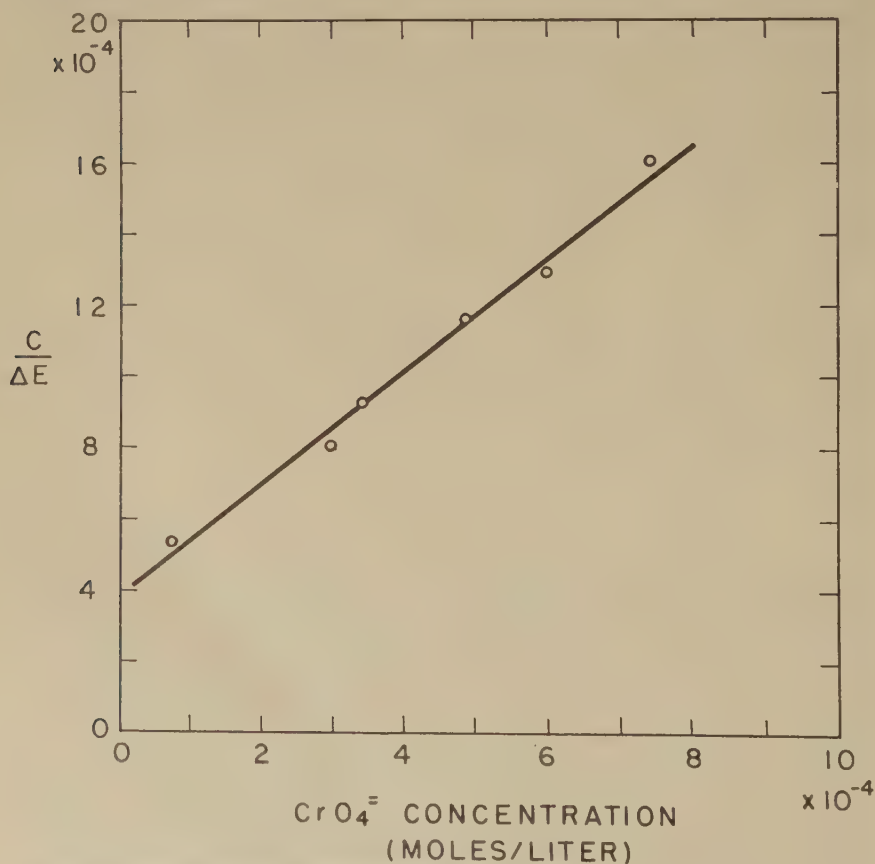


FIGURE 3. Langmuir adsorption plot for electrolytic iron in distilled water containing $\text{CrO}_4^{=}$ inhibitor. $T = 25.0 \pm 0.2^\circ\text{C}$.

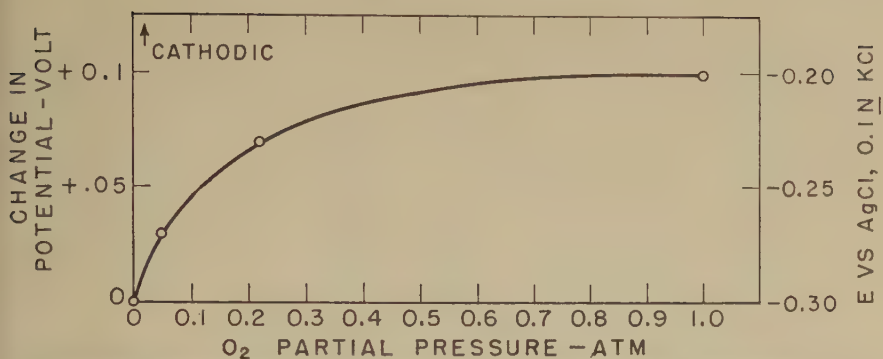


FIGURE 4. Effect of varying O_2 partial pressure on the electrode potential of 18-8 stainless steel in 4 per cent NaCl, 0.3 M NaOH. $T = 25.0 \pm 0.2^\circ C$.

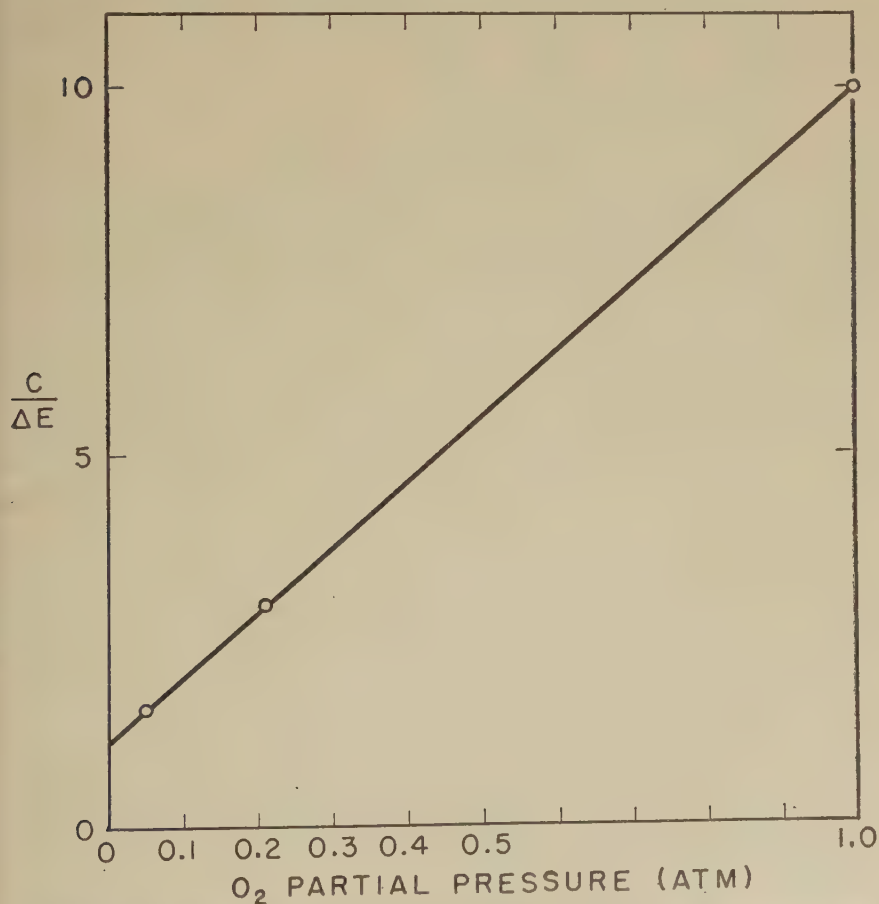


FIGURE 5. Langmuir adsorption plot for 18-8 stainless steel in 4 per cent NaCl with 0.3 M NaOH, varying O_2 partial pressure. $T = 25.0 \pm 0.2^\circ C$.

Passivity

The relation of passivity to chemisorption reasonably explains why carbon monoxide, which chemisorbs on several metals, is a successful passivator for 18-8 stainless steels immersed in hydrochloric acid.¹⁴ Any mechanism based on a protective oxide film is less plausible. Similarly, specific organic inhibitors will adsorb, on the surface of iron, the corresponding potential changes obeying the Langmuir adsorption isotherm as in the case of chromates on iron.¹⁵ The affinities are not so great as for chromates, so that the mechanism of protection is probably a combination of (1) physical protection by a close-packed organic molecular layer and (2) satisfaction of surface affinities.

This proposed mechanism of passivity leads to the conclusion that partial passivity can also result from a partial monolayer, if such a layer in some degree satisfies surface affinities. Also, since a second layer over a first chemisorbed layer may possibly attach itself through holes in the first layer and bond itself to the metal more strongly than a typical physically adsorbed layer, a degree of increased passivity will result from the second layer, but relatively little effect is expected from a third or from subsequent layers.

Factors Determining Chemisorption and Passivity

The requirements for passivity* as just described are (1) a high affinity of the environment for the metal surface; (2) a high activation energy for the reaction of adsorbate with metal, or correspondingly slow formation of a surface stoichiometric compound. These conditions correspond to those necessary for chemisorption. The surface compound may ultimately act as a diffusion barrier and thereby still further reduce the reaction rate, but the primary cause of passivity is the chemisorbed layer which maintains itself at the metal surface in the absence or presence of a compound.

Since a high activation energy for reaction is essential, it is obvious that chlorides or chlorine with zero or low activation energy of reaction, even though having the requisite metal affinity, do not passivate either iron or the stainless steels. In fact, they destroy passivity by competing with and displacing adsorbed oxygen, which is one of the successful passivators.

One factor within a metal favorably disposing it to formation of chemisorbed films is a high heat of sublimation, since atoms of the metal are then more reluctant to leave the crystal lattice. Another factor is a low work function or easy surrender of an electron within the metal to the environment. Metals of this kind are usually the transition metals of the Periodic Table characterized by unfilled *d* electron energy bands. They proceed from scandium through nickel in the first sequence, from yttrium through palladium in the second sequence, and from hafnium through platinum in the third. A simple expression can be set up illustrating the relation of sublimation energy and work function to chemisorption and passivity. If an electron escapes from a metal, as measured by its work function, more readily than can a metal atom, as measured by its sublimation energy, such a metal favors chemisorption of substances such

* Passivity as defined in Definition 1 of *The Corrosion Handbook*, p. 21. Uhlig, H. H., Ed. 1948. Wiley. New York, N. Y.

TABLE 1
RATIOS OF WORK FUNCTION TO HEAT OF SUBLIMATION FOR VARIOUS METALS

	Heat of sublimation, ΔH_s , 25°C, electron-volts	Work function, electron-volts	Ratio, work function $/\Delta H_s$
K	0.93	2.24	2.4
Na	1.12	2.37	2.1
Ca	1.85	3.09	1.7
Mg	1.56	3.60	2.3
Zn	1.35	3.7-4.2	2.7-3.1
Hg	0.63	4.50	7.1
Pb	2.01	3.8	1.9
Al	2.92	3.6-4.4	1.2-1.5
Ag	3.0	3.9-4.7	1.3-1.6
Cu	3.54	4.80	1.36
Mn	3.02	3.76	1.25
Pt	5.40	6.3	1.17
Ni	4.26	4.1-5.0	0.96-1.2
Fe	4.20	4.2-4.7	1.0-1.1
Cr	3.87	4.35	1.1
W	8.8	4.52	0.51
Mo	6.75	4.12	0.61
C	8.6	4.82	0.56

as oxygen. Oxygen tends to absorb electrons, and the volatility of the metal-oxygen ion complex is sufficiently low so that the metal atom remains in its lattice. On the other hand, if the atom escapes more readily than does the electron, the metal preferably forms an oxide, chemisorption does not occur, and the metal is not passive. Should we express both sublimation energy ΔH_s and work function ϕ in electron volts, then if $\phi/\Delta H_s$ is less than one, conditions are favorable to passivity and chemisorption, whereas if the ratio is greater than one we have oxidation and lack of passivity. A list of such ratios is shown in TABLE 1 illustrating that nontransition metals like sodium and copper oxidize readily and are not passive (as we presently define passivity), whereas transition metals like chromium and nickel are more nearly in the second category, tending to become passive.

Effect of Electron Configuration

Rates of chemical reaction frequently depend on catalysts capable of chemisorbing some component of the reacting species. Hence, the transition metals with unfilled d bands comprise some of the most active catalysts. An example in point is the catalytic conversion of parahydrogen, stable at low temperatures, to an equilibrium mixture of para- and orthohydrogen at room temperature. A layer of chemisorbed hydrogen atoms on a suitable catalyst such as palladium exchanges with molecular hydrogen much more rapidly than does molecular ortho- with molecular parahydrogen; hence, equilibrium is accelerated by presence of the metal. Palladium has 0.6 electron vacancy per atom in the d band of energy levels, as determined by magnetic data. Gold, on the other hand, is a nontransition metal with a filled d band and is correspondingly a poor catalyst. The two metals form solid solution alloys with electrons of alloyed gold

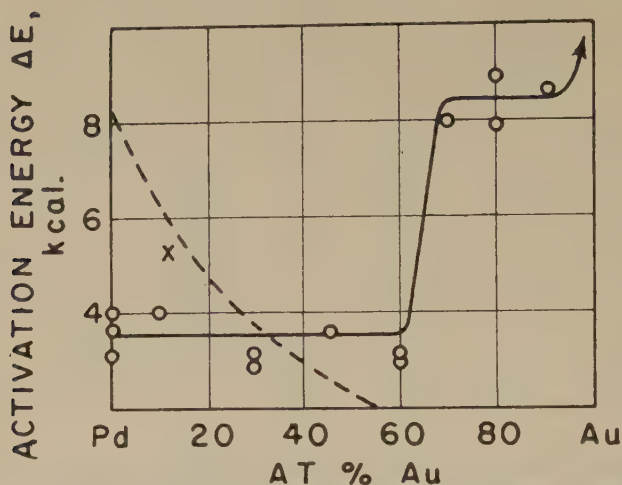


FIGURE 6. Activation energy for ortho-para hydrogen conversion on Pd-Au alloys. Broken line X denotes the paramagnetic susceptibility in arbitrary units, and indicates a filled d band above approximately 60 At. per cent Au. (Couper and Eley)

tending to fill the partially empty d band of palladium. So long as the d band of the alloy is partially empty, the alloy is a relatively good catalyst, but when the d band fills, which occurs at 60 atom per cent gold or higher, catalytic properties deteriorate. This correspondence of electron configuration with catalytic efficiency was suggested by Couper and Eley¹⁶ and confirmed by them experimentally. FIGURE 6 illustrates the trend of activation energy ΔE with per cent Au.* Above 40 atom per cent Pd, in other words, the d band is partially empty and chemisorption of hydrogen occurs accompanied by a low activation energy for reaction. Below this concentration, only physical adsorption is expected, and the alloys are no better catalytically than gold. A further fact supporting the importance of an unfilled d band to this reaction is the fact that if hydrogen is dissolved in palladium, the electrons of dissociated hydrogen atoms fill the d band of energy levels in the same manner as do electrons from gold, thereby diminishing catalytic activity.

Since passivity, as discussed above, also appears to depend on chemisorption, similar relations of passivity to electron configuration of metal or alloy might be expected. This is precisely what is found for several alloy systems, and provides a reasonable explanation for the well-known critical compositions for passive or corrosion resistant alloys, similar to the critical composition of Au-Pd catalyst. It also explains why such alloys are largely made up of the transition metals. These are precisely the metals which favor chemisorption rather than surface compound formation.

One of the best examples illustrating the relation between electron configuration and passivity is provided by the copper-nickel system. Nickel is passive with 0.6 electron vacancy per atom in the d band, but copper has the full component of 10 d electrons making the d band full. In hot NaCl solutions,

* The A value of the Arrhenius equation is relatively constant with alloy composition.

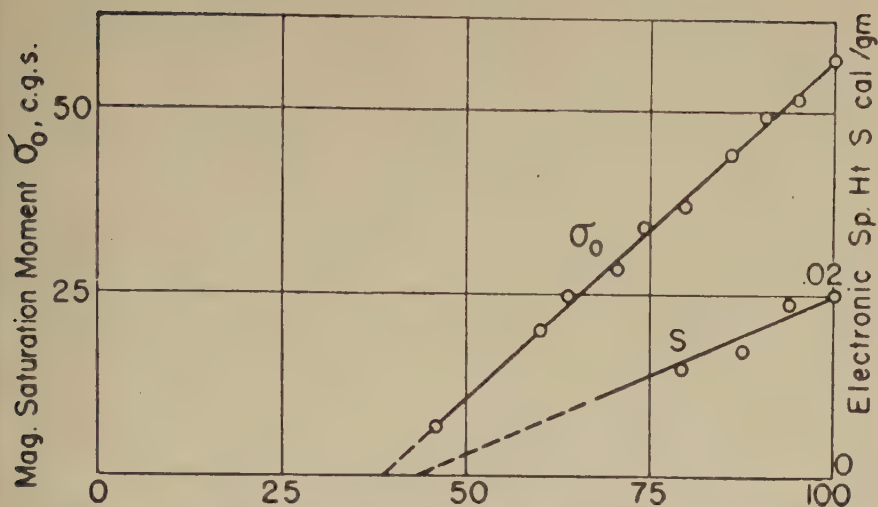


FIGURE 7. Variation of the magnetic saturation moment and electronic specific heat with composition of Cu-Ni alloys. The d band is filled above approximately 60 At. per cent Cu.

copper, which is not a passive metal, corrodes at a higher rate than does nickel. When nickel is alloyed with copper, the two metals form a solid solution system throughout the composition range, with one electron from each copper atom available to fill the d band of nickel. When the nickel content of the alloy reaches a certain level, electrons from copper are no longer sufficient to fill the band. Alloys above this critical nickel composition behave as transition metals and are passive, behaving more like nickel, in other words, than copper; whereas alloys below the critical composition behave more like copper. Magnetic and specific heat data¹⁷ provide information that the d band is partially empty above 40 atom per cent (38 wt. per cent) nickel (FIGURE 7) which correlates with corrosion data obtained recently by H. Keily and J. Iannicelli of the Massachusetts Institute of Technology Corrosion Laboratory. Their results are shown in FIGURE 8. The corrosion data confirm that a change in rate occurs at approximately the predicted composition, the alloys containing more than 30 to 40 wt. per cent nickel corroding at a lower rate in 3 per cent NaCl at 80° C. than alloys of lower nickel content. Similar conclusions are reached from data of Friend¹⁸ and LaQue¹⁹ on corrosion of copper-nickel alloys in sea water. Alloying in this instance presumably alters the ratio of $\phi/\Delta H_s$, high nickel alloys having a value less than unity and favoring chemisorption and passivity.

Parallel relations are found in other passive alloy systems, even though the physical details of electron configuration are not as well understood as for the examples cited above. The chromium-iron alloy system is probably most familiar. These alloys containing more than the critical composition of 12 to 15 per cent Cr compose the familiar stainless steels, whereas alloys of less than 12 per cent Cr are not passive and corrode much more rapidly. It is natural to assume that the d band of chromium tends to fill with electrons supplied by

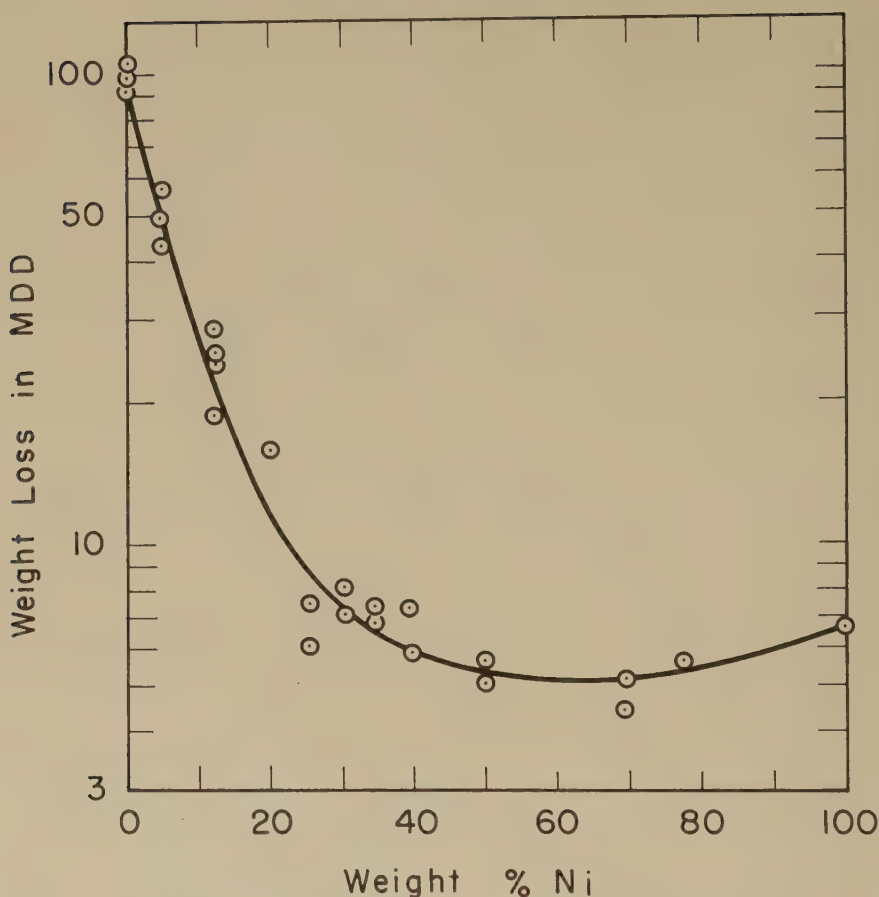


FIGURE 8. Corrosion rates of Cu-Ni alloys in aerated 3 per cent NaCl at 80°C., two-day exposure.

iron, since chromium is the element naturally passive in air or water solutions, whereas iron is not passive under similar circumstances.* The d band of chromium, therefore, can be assumed to be effectively filled below the critical chromium concentration for passivity, but unfilled above. Some properties of Cr-Fe alloys bear out an electronic change at the 12 per cent Cr composition as, for example, electric resistivity.²⁰ By making certain assumptions concerning demotion of s electrons into vacant d levels of the metal when the metal atoms condense to a solid, and further assuming that demoted electrons of this kind are available to fill the d band of the alloy,²¹ a surprisingly satisfactory correlation of predicted and observed passive compositions can be found. The observed critical passive compositions are those observed, for

* Oxygen readily chemisorbs on both chromium and on iron, but in the presence of water, the activation energy for reaction becomes practically zero in the case of iron, and remains at an appreciable level in the case of chromium.

TABLE 2
ELECTRON CONFIGURATION AND CRITICAL ALLOY COMPOSITIONS FOR PASSIVITY

Alloy	Atomic configuration	Critical composition for passivity (In dil. H ₂ SO ₄ unless otherwise indicated)			
		Predicted		Observed	
		At. %	Wt. %	At. %	Wt. %
Cr	3d ⁵ 4s	16.7% Cr	15.7% Cr	13-16% Cr	12-15% Cr*
Fe	3d ⁶ 4s ²				
Ni	3d ⁸ 4s ²	40% Ni	38% Ni	32-42% Ni	30-40% Ni†
Cu	3d ¹⁰ 4s				
Ni	3d ⁸ 4s ²	33½% Ni	34% Ni	39% Ni	40% Ni
Fe	3d ⁶ 4s ²				
Co	3d ⁷ 4s ²	25% Co	26% Co	20-25% Co	21-26% Co
Fe	3d ⁶ 4s ²				
Mo	4d ⁵ 5s	9.1% Mo	14.7% Mo	9% Mo	15% Mo‡
Ni	3d ⁸ 4s ²				
Mo	4d ⁵ 5s	42.9% Mo	55% Mo	52% Mo	65% Mo
Fe	3d ⁶ 4s ²				

* Dilute sulfates.

† 4% NaCl at 80°C. In H₂SO₄, Cu⁺⁺ acts as depolarizer.

‡ HCl.

example, in sulfuric acid. The critical compositions may shift in either direction in other media, depending on factors such as oxidizing capacity and concentration of the environment, and the nature of ions present and their tendency to adsorb. Correlation of this kind is summarized in TABLE 2.

Conclusions

- (1) Unsatisfied forces at the surface of a metal lead to:
 - (a) a state of tension of surface atoms and compression of atoms underneath; and
 - (b) the formation of adsorbed films physically or chemically bonded to the surface.
- (2) Chemisorbed layers account for:
 - (a) catalytic activity of many metals and alloys; and
 - (b) passivity as observed in chromium, nickel, stainless steels, and in transition metals or their alloys in general.

Conditions necessary for passivity are:

 - (a) high affinity of the environment for the metal; and
 - (b) high activation energy for reaction to form stoichiometric compounds.
- (3) Chemisorption or, alternatively, reaction of metal with the environment, depends on:
 - (a) sublimation energy of the metal, high values favoring chemisorption and passivity; and

- (b) work function of the metal, low values favoring chemisorption and passivity.
- (4) These properties are related to electron configuration of the metal. The unfilled *d* band of transition metals correlates with their passive properties. Any situation which, in effect, fills the *d* band destroys passivity, accounting for a critical composition range in passive alloys where one or more components act as electron donors.

References

1. PAULING, L. 1947. J. Am. Chem. Soc. **69**: 542.
2. UHLIG, H. H. 1952. Metal interfaces. Am. Soc. Metals. Cleveland, Ohio. : 313.
3. PAULING, L. 1948. Nature of the Chemical Bond. Cornell Univ. Press. Ithaca, N. Y.
4. HOFMANN, U. & D. WILM. 1936. Z. Elektrochem. **42**: 504.
5. BENEDICKS, C. 1948. Pittsburgh International Conference on Surface Reactions. : 196.
6. BUTTNER, F. H., E. R. FUNK, & U. UDIN. 1952. J. Phys. Chem. **56**: 657.
7. LANGMUIR, I. 1916. J. Am. Chem. Soc. **38**: 2221.
8. LENNARD-JONES, J. E. 1932. Trans. Faraday Soc. **28**: 333.
9. ELEY, D. D. 1949. Quart. Rev. Chem. Soc. (London, England.) III. : 209.
10. LANGMUIR, I. 1916. Trans. Electrochem. Soc. **29**: 260; J. Chem. Soc. (London). **1940**: 518.
11. LANGMUIR, I. 1932. J. Am. Chem. Soc. **54**: 2798.
12. UHLIG, H. H. & A. GEARY. 1954. J. Electrochem. Soc. **101**: 215.
13. ROBERTSON, W. D. 1951. J. Electrochem. Soc. **98**: 94.
14. UHLIG, H. H. 1940. Ind. Eng. Chem. **32**: 1490.
15. N. HACKERMAN & J. SUDBURY. 1950. J. Electrochem. Soc. **97**: 109.
16. COUPER, A. & D. ELEY. 1950. Disc. Faraday Soc. : 172.
17. MOTT, N. & H. JONES. 1936. Theory of Properties of Metals and Alloys : 199. Oxford Univ. Press. London, England.
18. FRIEND, J. 1932. J. Inst. Metals. **48**: 109.
19. LAQUE, F. L. 1941. Am. Soc. Naval Engrs. **53**: 29.
20. ADCOCK, F. 1931. J. Iron Steel Inst. **124**: 99.
21. UHLIG, H. H. To be published.

OXIDE FILM COMPOSITION STUDIES

By Thor N. Rhodin, Jr.

Engineering Research Laboratory, E. I. du Pont de Nemours and Company, Wilmington, Del.

Introduction

The unusual surface property of stainless and related alloys of resisting corrosion in an environment where thermodynamic considerations predict reaction has been known for 40 years.¹ The basic mechanisms associated with this phenomenon, passivity, remain unestablished after considerable study.² The purpose of this present work is not to substantiate any one of several available theories or to compare their relative merits, but to provide basic information on oxide film properties by which the proposed theories may be more effectively evaluated. The generalized film theory² emphasizes the importance of protective films in defining passivity. Considerable evidence indicates that thin oxide films in the thickness range of ten monolayers are closely connected with the development of passivity,^{3, 4, 5, 6, 7} although it is true that elucidation of the mechanisms of passivity must eventually correlate data on corrosion rates and electrode potentials as well as oxide film properties. This theory may be interpreted to include films other than oxides, although most of the study deals with oxide films. It is neither essential that certain oxide films be the unique cause of passivity nor that all oxide films on stainless alloys be equally effective.⁸

Studies of films isolated from passive stainless alloys present direct evidence distinguishing them from other oxide films and point to their presence as being connected with unique corrosion behavior. Many types of oxide films have been confused by investigators in efforts to evaluate the generalized film theory. Compositional data on passive films in sufficient detail will contribute to a better understanding of those particular oxide film properties that are critically associated with passivity. Earlier studies⁷ in this laboratory and elsewhere⁴ have suggested that passive films may have unique compositional properties. It is the purpose of this work to test this hypothesis and to point out where compositional film properties correlate with structural film properties in promoting passivity.

Procedure

The chemical and physical properties peculiar to passive films render them difficult to handle, easily contaminated, extremely sensitive to treatment and environment, and markedly affected by trace impurities. A detailed description of the sampling and analytical methods developed to overcome these difficulties is reserved for another communication. Adequate description of the procedures will be presented in this paper, however, to provide a background against which the data can be analyzed and the conclusions evaluated.

In the present study, surface films on stainless alloy Types 304, 316, 347, 302B, and on 14.5 per cent silicon-iron alloy were investigated. The chemical composition of these alloys is listed in TABLE 1. All specimens were annealed

TABLE 1
 ALLOY COMPOSITIONS

Alloy	Fe	Cr	Ni	Si	Mo	C
	%	%	%	%	%	%
304	71.7	18.3	8.70	0.6	—	0.07
347	68.0	18.2	10.80	0.6	0.2	0.04
316	64.7	17.6	13.00	0.6	2.3	0.08
302B	70.0	17.3	8.60	2.7	—	0.14
446	71.7	26.3	0.08	0.5	—	0.05
Silicon-Iron	84.0	—	—	14.5	—	0.85

foil given a bright pickle with 10 per cent HNO_3 -2 per cent HF -1 per cent HCl solution at 80°C ., except for the silicon-iron alloy, which were small plates pickled in boiling 20 per cent sulfuric acid. The specimens were washed in distilled, deionized, low-conductivity water and most of them given an immersion passivation in 5 per cent HNO_3 -0.5 per cent $\text{K}_2\text{Cr}_2\text{O}_7$ solution for 30 minutes at 60°C ., except for the silicon-iron alloy, which was immersed in 30 per cent boiling HNO_3 for 10 minutes. Not every sample was given the complete treatment; some were removed at the end of each stage for surface film analysis.

Films formed on stainless alloys by these treatments have shown excellent corrosion resistance to halide ion pitting attack in 10 per cent FeCl_3 solution at room temperature.⁹ It is recognized that no passivation treatment confers lasting protection in media corrosive to unpassivated alloys and that the passive films of this study may not be equally effective in other corrosive media. The passivating conditions used were chosen as being the conditions most likely to yield information of general significance. The importance of extending compositional studies to surface films produced after long-time immersion in corrosive media is also recognized. It is hoped that this study will stimulate research on film compositional properties characteristic of these more complicated systems.

Film Stripping Procedures

Sampling a surface to yield representative and uncontaminated material proved to be a formidable problem. Experiments to evaluate methods for stripping very thin films were made, using chemical, electrolytic, and mechanical techniques. Two satisfactory sampling methods were finally developed and evaluated, each having its own merits and limitations: (1) a chemical method using 2 per cent bromine-anhydrous methanol solution, and (2) a mechanical method using a high-purity pressure-sensitive synthetic rubber adhesive on "Mylar" tape. The bromine-methanol reagent, first used by Nielsen and Mahla,⁷ was preferred because it proved to be most effective for controlled and reproducible film stripping. An anhydrous methanol-2 per cent bromine solution in an atmosphere of pure nitrogen was found to strip film quantitatively when used in a specially designed all-glass stripping apparatus* illustrated in FIGURE 1, similar to the type first used by Vernon, Wormwell, and

* Designed with the collaboration of H. W. McGonigal.

Nurse.⁴ The sample is placed in position and the system is flushed with pure nitrogen. After pickling, passivating, and washing *in situ*, the sample is immersed in methanol, and a low-ash Whatman No. 44 filter paper is placed on the filter disk. The sample is immersed in 2 per cent bromine-methanol which undermines the surface film. Gentle bubbling of nitrogen along the surface disengages the film fragments, which are floated over to the collector by siphoning action through the gooseneck connector. The film sample is thoroughly washed by passing a relatively large volume of methanol through the filter disk. The receiver is disassembled, the filter disk is removed, and the filter paper carefully transferred to a 2-ml. platinum crucible for ignition, weighing, *etc.* The stripping operation is conducted in an enclosed controlled-atmosphere hood, and blank runs are made at regular intervals as a precaution against unknown contamination of one sample by another or by the stripping apparatus itself. Development of this technique made possible the convenient isolation of oxide films weighing as little as 10γ ($1\gamma = 10^{-6}\text{g.}$) in a continuous operation with minimum sample loss, and no significant contamination or air exposure.

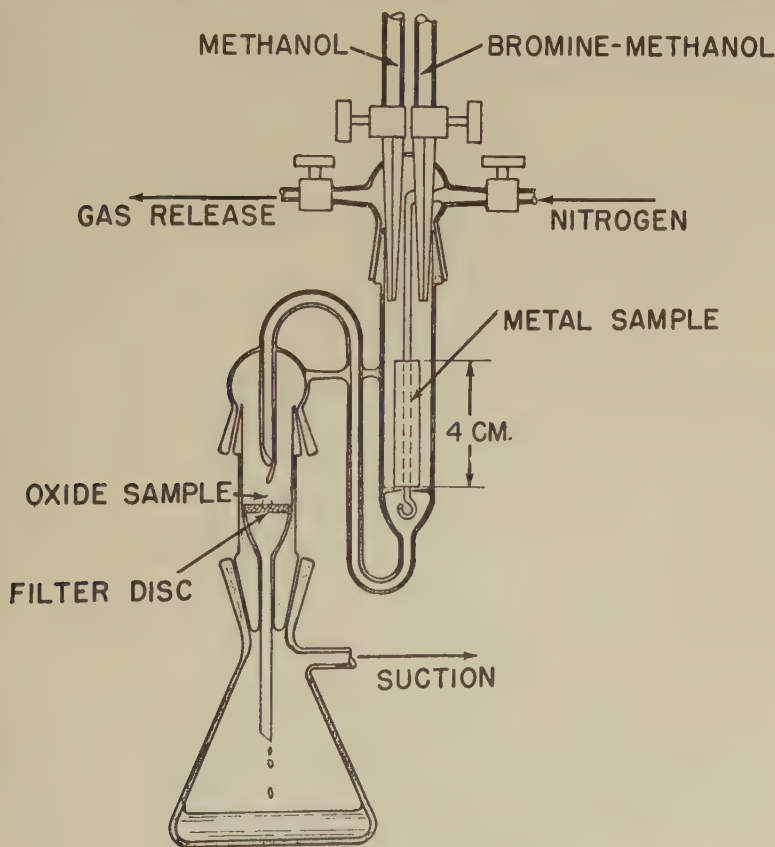


FIGURE 1. Film stripping apparatus.

In the absence of air or moisture, successive stripping was not possible, confirming previous reports⁹ that the stripping solution did not produce the film that is isolated. That the film formed during passivation is the one isolated with the stripping reagent has also been indicated by Simnad, using radioactive chromium.¹¹ Introduction of air into the apparatus onto a freshly stripped sample did result in a new surface film which could be readily collected. Data on the average thickness of passive films have been published^{12, 13, 14, 15} but uncertainty in the values used for surface roughness, film density and composition make their interpretation of questionable usefulness. It is interesting to note, however, that quantitative collection of passive film from samples of known surface area yielded 0.5 to 1.0 γ of film sample/sq.cm. of alloy surface. Assuming the metals are present in their highest oxidation states, calculations based on the film compositional data indicate from 0.15 to 0.30 γ of O₂/sq.cm. of passivated alloy surface.

Sources and cause of silica contamination reported by Nielsen and Mahla⁷ and by Fontana¹⁶ were evaluated with special reference to the stripping techniques used. Contamination during bromine-methanol sampling by silica inclusions dissolved out of the metal during stripping did occur. Typical data on the relative amounts involved is illustrated in FIGURE 2. By accurate evaluation of the effect, optimum stripping conditions involving 1 per cent or less dissolution of base metal were developed and suitably small corrections were applied. In view of possible silica contamination by the stripping apparatus itself, blank runs during the course of the study indicated that pick-up of silica during stripping was small and never exceeded 15 per cent of the total silicon determined in a film sample. Silicon contamination from glass apparatus containing water solutions was a more serious difficulty than the relatively small error introduced by the stripper. It was eliminated only when the samples were manipulated completely free of glass apparatus after stripping. Reliable silicon data were obtainable only when all glass equipment was replaced with microbeakers, volumetric flasks, and adsorption cells of polyethylene and polytetrafluorethylene. Extensive calibration studies were made on silicon pick-up from glass apparatus by water solutions in terms of pH, dissolved metals, and surface condition of the glass. Even new glass was found to introduce as much as 0.01 γ /(sq.cm.)/(day) of silicon into neutral distilled deionized water with an initial resistivity of 13 megohms. Freshly boiled deionized water contained in polyethylene in comparison retained a low silicon blank of less than 0.01 γ /ml. indefinitely. After quantitative and systematic study of known sources of silicon contamination, it was finally concluded that the presence of appreciable amounts of silicon in the passive films was definitely established.

Exploratory studies of stripped oxide films indicated composition was not affected significantly by prolonged immersion in the stripping solution, indicating the stability of the oxides in the dilute bromine-methanol solution. Any unreacted metal incorporated in, or adhering to, the film after stripping would, of course, be dissolved prior to analysis. It is believed the stripping solution cannot promote oxide film formation under the stripping conditions used and

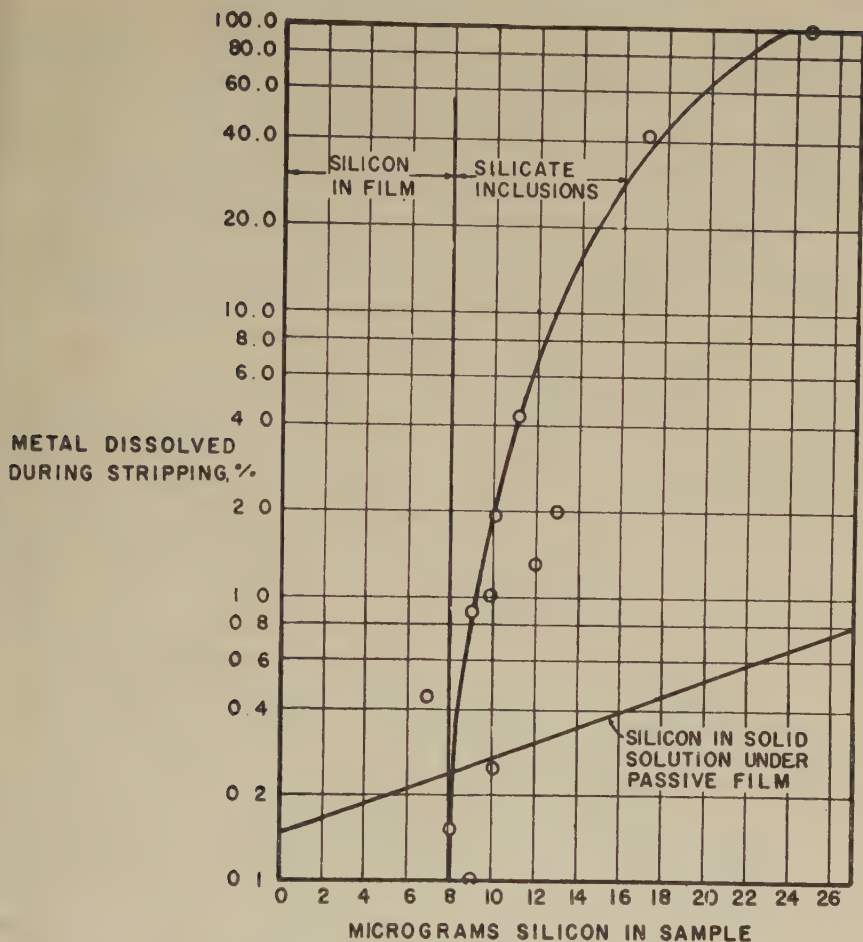


FIGURE 2. Silicate contamination of passive film S.S. Type 304.

that the stripping agent itself does not seriously influence the compositional character of the film sample.^{4, 7, 10}

Finally, the reliability of the chemical stripping technique was evaluated by comparing it with data obtained, using a completely independent stripping method. This method was successful mainly because of the unusual high purity of "Mylar" tape and rubber adhesive* and the aggressiveness of the adhesive in forming a strong bond with both tape and oxide film.

The adhesive side of a "Mylar" tape was pressed strongly against a freshly passivated surface with sufficient force to fragment some of the film. Upon stripping the tape from the surface, fragments of film adhering to the adhesive were also removed. The adhesive was dissolved in a Slotta¹⁷ extraction ap-

* Spectroscopic analysis: Fe, 300 p.p.m.; Si, 15 p.p.m.; Ni, negligible.

paratus, and the film fragments were collected on a microfilter paper supported on a filter disk. The size of samples obtained was low (5–10 γ), but they were reproducible for a given surface. Except for the necessity of eliminating samples containing large inclusions, and the requirement that blank corrections be checked on each run, this method proved remarkably effective. It provided an excellent check on the reliability of the modified bromine-methanol stripping method. A third stripping method was also used to obtain large sections of passive film (2–4 sq.cm.). It was a modified bromine-methanol technique* in which the film was supported with a "Formvar" backing, placed on a "Mylar" support, and the metal backing dissolved. The method was effective for isolating large sections of film for study by optical and electron metallography.

Microanalysis of Films

The general procedure employed was a complicated one in which the samples were stripped, washed, ignited, weighed, fused, leached, and finally analyzed spectrophotometrically. The flow chart in FIGURE 3 summarizes the main operations. Approximately 100 separate samples were analyzed. Of these, 30 samples were eliminated from consideration because of lack of mass balance before and after analysis, or because of known contamination during handling, or whenever the sum of the blank corrections became too high. Since most of the blank corrections totaled less than 20 per cent of the final value determined for a single component, any value of a single blank in excess of this percentage was arbitrarily taken as being too high. Commercially available reagents of the highest purity were used. Rather than run hazards of contamination during additional purification, few of the reagents were repurified but analyzed in blank runs using the standard analytical procedures. Purities of standard reagents were found to vary considerably, but eventually a group of acceptably pure materials was collected. Suitable manipulative procedures were developed which avoided the need for using reagents in inordinately large quantities relative to the size of the samples. In this manner, serious contamination of the samples by trace impurities in the reagents was greatly minimized, and in no case did a cumulative blank correction for all the impurities introduced by the reagents during a complete analysis amount to more than 20 per cent of the element being analyzed. It was concluded during the study that reliable data were obtainable only by frequent revaluation of blank corrections and restandardization of reference solutions. Blank corrections were applied to all data. Ignition of film samples at 1000° C. for one hour yielded samples weighing 10 to 100 γ which could be weighed to $\pm 1\gamma$ on an Ainsworth microbalance. To achieve this precision, it was necessary to make a special installation of the instrument to eliminate interference from building vibration, dust, air currents, humidity, and temperature changes. A two-stage microfusion technique was developed to convert the refractory oxides into water-soluble form. The first fusion in Na_2CO_3 at 850° C. quantitatively fused the oxides of silicon, chromium, molybdenum, and niobium. After leaching in deionized water and filtering the residue, the oxides of iron and nickel were fused in $\text{Na}_2\text{S}_2\text{O}_7$ at 650° C. and leached.

* Developed with the collaboration of R. H. Fisher.

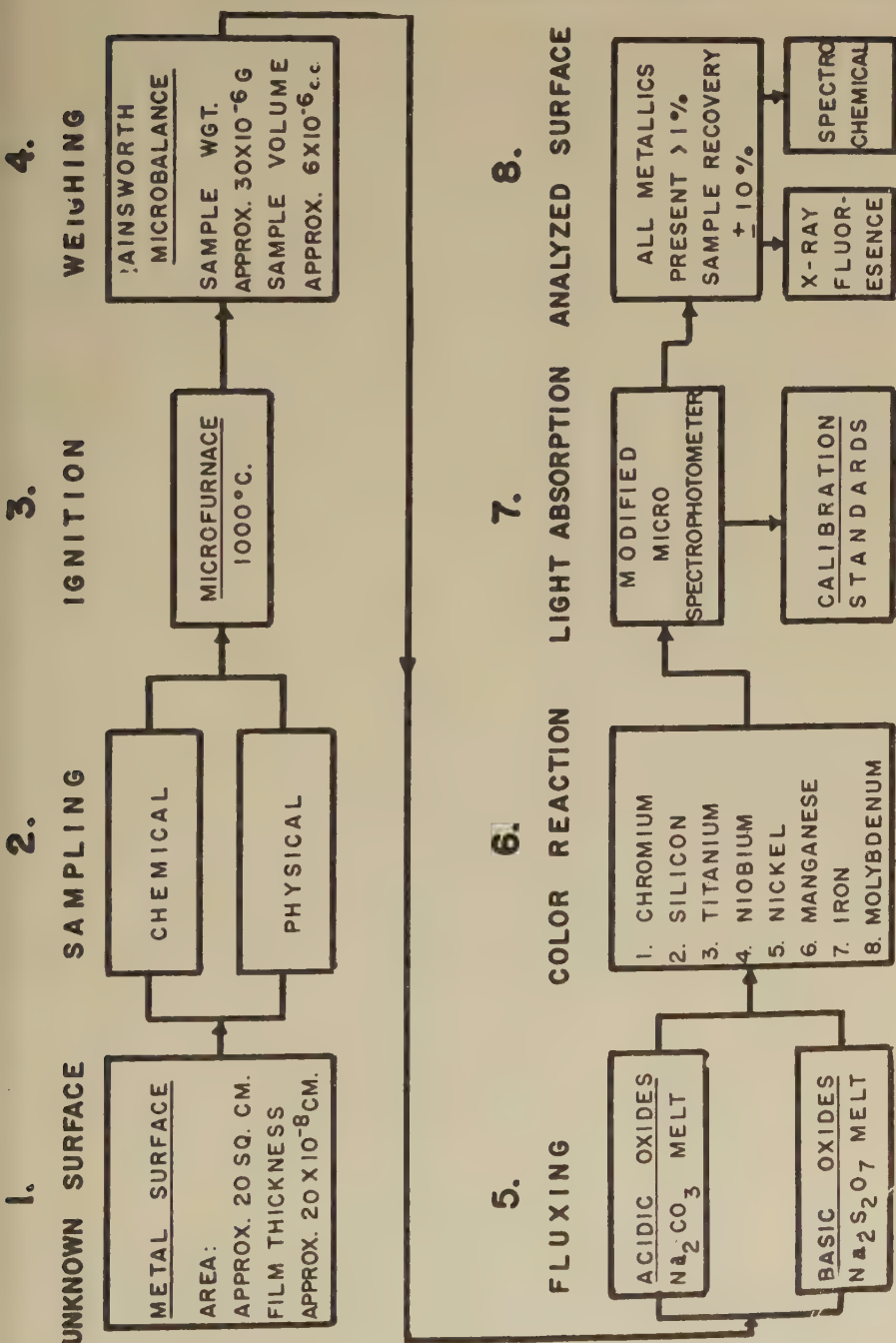


FIGURE 3. Simplified flow chart of microanalytical method.

The colorimetric procedures used for the microanalysis of iron^{18, 19, 20} nickel,²¹ chromium,^{22, 23} manganese,^{24, 25} and silicon^{26, 27} were refined modifications of the macroscopic analytical procedures described in the literature. The difficulties in scaling down were usually those of contamination, micromanipulations, chemical interferences, or inadequate sensitivity. Contamination was minimized by controlling reagent purity and sample manipulation. Micromanipulations were largely a problem of developing suitable skills and techniques. Chemical interferences were avoided by using microseparations in the preliminary stage by fusion-filtration or, for particular elements, by extraction at the color development state. Sensitivity was improved by using very small liquid volumes (0.1–5.0 ml.) and high effective concentrations of the spectrophotometric samples. A special microabsorption cell²⁸ of polytetrafluorethylene with a 7.5-cm. light path and a 4-mm. internal diameter was also used, yielding an increase in sensitivity of 10× over the standard cell with 1-cm. path length and 20-mm. diameter. Careful alignment was critical for the microcells, requiring a cell holder with three-directional micrometer screw adjustments. The scaling down of colorimetric methods for the analysis of niobium^{29, 30} and molybdenum^{31, 32} was complicated considerably by the lack of reliable macroscopic methods, but was finally achieved with moderately good sensitivity. The general characteristics of all the microcolorimetric methods used are summarized in TABLE 2 with references to the related macroscopic methods, together with the sensitivities achieved for pure separate standards on the microanalytical scale. The details of the microanalytical procedures resulting from the refinement of the macroscopic procedures are omitted because they are extensive and not essential to the interpretation of the data.

Standardization curves were prepared for all the metallic elements and the procedures were cross-checked against mixed oxide standards prepared in the laboratory from oxides of known purity. By basing the standardization curves on mixtures of oxides, corrections were made for loss of sensitivity and other errors caused by interferences, blank corrections, and micromanipulations characteristic of the actual analytical procedure. This procedure was essential in obtaining reliable analytical data. The reliability of the procedures (TABLE 3) indicates that the actual sensitivities (Column 3) achieved in practice for known oxide mixtures are approximately 10× lower than the sensitivities

TABLE 2
CHARACTERISTICS OF THE MICROCOLORIMETRIC METHODS

Metal	Method	Wavelength, Å	Absorption coefficient	Analytical sensitivity (γ/ml./cm.)
Cr	Diphenylcarbazide	5400	720	0.002
Si	Molybdenum Blue	8200	612	0.006
Fe	o-Phenanthroline	5040	195	0.015
Mo	Thiocyanate	4750	98	0.024
Nb	Thiocyanate	3850	180	0.015
Ni	Dithiocarbamate	3850	102	0.040
Mn	Permanganate	5250	45	0.100

TABLE 3
RELIABILITY OF THE ANALYTICAL PROCEDURES

Metal	Analytical sensitivity (γ /ml./cm.)	Actual sensitivity (γ /ml./cm.)	Experimental variation (γ /ml./cm.)
Cr	0.002	0.01	± 0.20
Si	0.006	0.05	0.10
Fe	0.015	0.05	0.26
Mo	0.024	0.08	0.22
Nb	0.015	0.10	0.15
Ni	0.040	0.15	0.18
Mn	0.100	0.20	0.20

achieved for pure separate standards (Column 2). The experimental variation (Column 4) refers to the actual uncertainty observed for typical film samples isolated from different areas of the same alloy under identical conditions. Since, with the exception of the manganese procedure, these variations are, in general, substantially higher than the precision of the analytical or sampling procedures would predict, they are interpreted to reflect actual variations in composition over the surface of the metal. An independent study of composition of evaporated stainless alloy films and the passive films formed on them was consistent with this interpretation. Detailed discussion of that study will not be included here. It is interesting to note, however, that the evaporated alloys yielded passive film compositions, having substantially lower standard deviations. This finding is consistent with a greater uniformity of surface film composition for the evaporated alloys. The average values for the various film components, however, on both types of samples were very similar. The passive films on the evaporated samples were analyzed both microcolorimetrically and by a modified X-ray fluorescence method developed in collaboration with the North American Philips Company. A comparison of the microanalytical methods for data on iron, nickel, chromium, and silicon in passive film on Type 304 stainless alloy is illustrated in FIGURE 4. The data obtained by X-ray fluorescence are included for comparison (it was not possible at that time to analyze for silicon by this method). The agreement is good when the complexities of sampling, microanalysis, and uncontrolled surface variations characteristic of commercial alloys are considered. The silicon data obtained by the bromine-methanol method are high by an amount corresponding to the predicted silica contamination. All other silicon data reported are corrected for this effect.

A Beckman DUV quartz spectrophotometer with photomultiplier attachment and fitted with special microcells and cell compartment assemblies was used for all the microcolorimetric data.

Results and Discussion

Data on passive films formed on various typical stainless alloys are in agreement with predictions that film composition is unique for each alloy and related to, but not identical with, the bulk alloy composition. Film compositions were also found to depend on pretreatment of the surface prior to passivation.

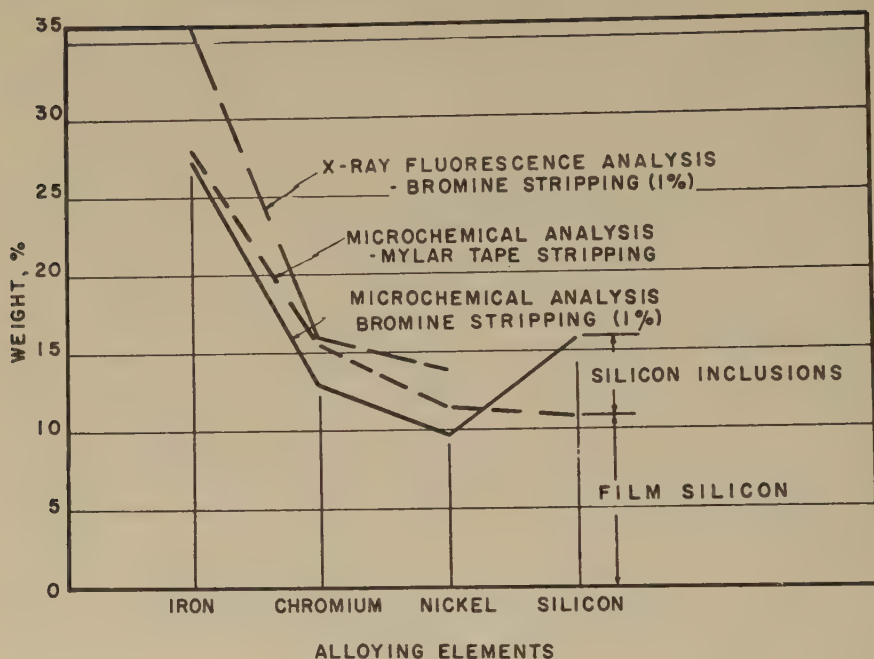


FIGURE 4. Comparison of microanalytical methods.

(1) Surface films for Type 304 stainless

(a) *Compositional characteristics:* Typical results on five samples of passive film isolated from stainless steel Type 304 prepared as described are summarized in TABLE 4. The compositional film data for stainless alloy Type 304 and all other films are expressed as per cent metal referred to the sample weight after ignition. The sample prior to ignition contained approximately 30 per cent water and after ignition also included approximately 30 per cent oxygen, presumably in the form of oxides. The assumption that the metals are present as oxides prior to ignition is based on the background of available information indicating the important role that oxidation processes play in their formation. Available electron diffraction data,^{7, 36} although not conclusive for passive films, indicates that film structures are composed of stoichiometric oxides for somewhat thicker films on stainless and related alloys.^{33, 34, 35} All metals are assumed to be present definitely as stoichiometric oxides after the ignition treatment. Unlike the composition of thicker oxide films,³³ known not to be critical in passivity, no appreciable enrichment in chromium nor depletion in nickel is observed for the passive films. This indicates, at least for this particular case, that passivity conditions are too complicated to be explained by the formation of a simple surface film enriched in Cr_2O_3 . It is also significant that the film associated with passivity shows an enrichment in silicon and a depletion in iron similar to that observed for thicker oxide films. It should be stressed that the

TABLE 4
TYPICAL DATA ON PASSIVE FILM
TYPE S.S. 304

Sample no.	Wgt.	Fe	Cr	Ni	Si	Recovery
	γ	%	%	%	%	%
1	55	26.8	14.3	16.0	9.8	119
2	89	33.3	10.5	7.5	8.6	102
3	74	35.0	6.8	5.8	10.8	105
4	90	23.3	15.6	8.4	8.8	96
5	36	18.0	18.0	11.1	11.9	106
Arith. mean.....	69	27.3	13.0	9.8	9.9	106
Stand. dev.....	—	7.1	4.4	4.0	2.0	9
Stoich. oxide.....	—	50.0	13.0	7.0	0.1	—

silicon data have been determined in glass-free apparatus, corrected for blank errors and for silica inclusions collected with the film. In addition, they are in good agreement with the silicon data for samples obtained by mechanical sampling which avoids contamination effects associated with chemical stripping. The data marked "stoichiometric" are characteristic of oxide film in which all metallic elements are present in proportion to the alloy composition.

Distinctions based on compositional data between thin oxide films associated with passivity and thicker oxide films associated with oxidation resistance at higher temperatures are indicated in TABLE 5. Comparison is made of film compositions of passive film on Type 304 and of thicker oxide films on Types 304 and 446. The latter are approximately 200 atom-layers thick and are produced by heating in air at 500° C. for one hour. Marked depletion in iron and nickel and enrichment in chromium and silicon is characteristic of the thicker films. In other words, as the oxide film thickens it appears that iron depletion increases and nickel depletion commences. Chromium enrichment initiates and silicon enrichment decreases concurrently. Other data of this study are consistent with these trends although systematic study of film composition with film growth was not made. The enrichment-depletion effects are accentuated for oxide film on Type 446. This is a straight-chromium ferritic stainless noted for its high-temperature scaling resistance. These com-

TABLE 5
COMPARISON OF PASSIVE FILM WITH THICK OXIDE FILM COMPOSITIONS—STAINLESS ALLOYS

	Passive† (S.S. 304)	Thick oxide (S.S. 304)	Thick oxide (S.S. 446)
	%	%	%
Iron.....	27	20	10
Chromium.....	14	35	50
Nickel.....	9	2	0
Silicon.....	10	8	6
Oxygen*.....	40	35	34

* Calculated by difference.

† Calculated for ignited film. Unignited film contains 35% water.

positional effects increase with temperature, indicating that a temperature-sensitive solid-state process such as solid diffusion, film recrystallization, or internal oxidation-reduction is important for thicker oxide films. In comparison, the compositional data on passive films are more characteristic of processes associated with selective dissolution followed by selective deposition of dissolved metals.

It is also interesting to note that the composition for the thick films on Type 304 are very similar to those reported ^{4, 5} for a similar stainless alloy polished with high-grade alumina where the film thicknesses are intermediate between passive and air-oxidized films. Vernon, Wormwell, and Nurse^{4, 5} suggest the chromium enrichment is associated with surface flow during polishing. Selective oxidation, controlled by restricted oxygen supply at the surface, is considered to be a controlling factor. Although their polished surfaces were considerably more passive than their stripped surfaces, the composition of the former suggests they are more closely related to thicker oxide films than those produced by chemical passivation.

An evaluation of the water content of passive films was made since films of this type are usually formed from aqueous media. Dehydration studies on passive films from Type 304 stripped with bromine-methanol showed a water content of 30 to 40 per cent by weight. Some of the dehydration data are summarized in FIGURE 5. Weight loss studies of passive films isolated from Type 304 stainless indicated water was present in a manner characteristic of a hydrous oxide in contrast to a hydrate or hydroxide. Samples ignited at 1000° C. showed negligible water uptake or loss, but samples dried at 125° C.

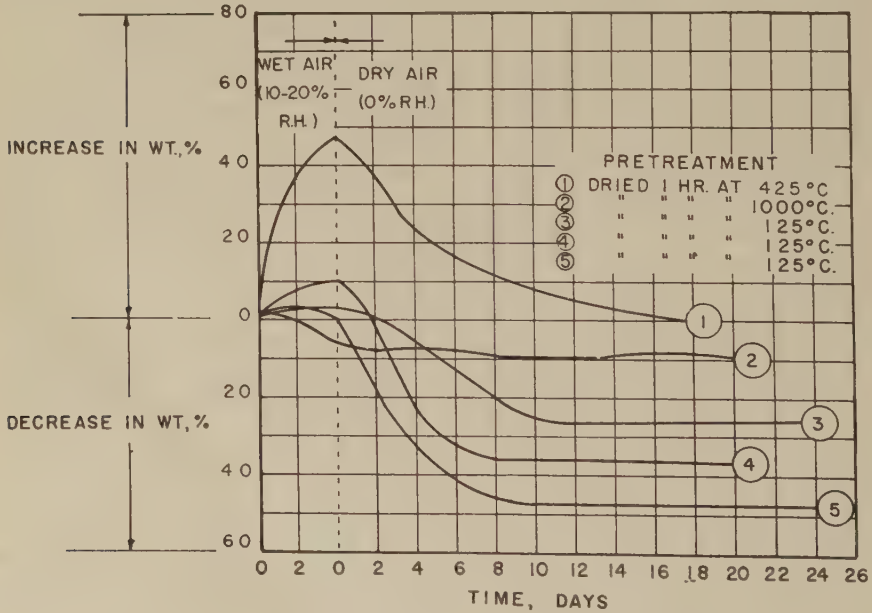


FIGURE 5. Water content of passive film S.S. Type 304 (Average weight 40 γ).

lost appreciable water on desiccation. Samples dried at intermediate temperatures adsorbed water in wet air and lost water in dry air corresponding to the quantity adsorbed. It is believed water is part of the passive film and is not removed during stripping but is removed on room-temperature desiccation or on heating above 400° C. Such a film sample adsorbs water on hydration, whereas neither ignited samples nor unignited samples adsorb water. It is suggested that the former have lost their adsorptive capacity by sintering and the latter are already saturated. The relatively high-water content of undried isolated passive films is characteristic of hydrous metal oxides with gel properties.³⁷

(b) *Structural characteristics*: Passive films isolated for analysis were also studied by light and electron metallography. A Philips electron microscope was used with facilities for making concurrent electron diffraction exposures on predefined areas of the film sample. The films were of the thin, transparent type substantially free of inclusions similar to those previously reported.⁷ They were also devoid of any observable unique surface texture at the highest magnification studied (25,000×). Their appearance in the electron microscope as well as their properties indicated by manipulations during microanalysis were characteristic of plastic, almost viscous, solids. These films differ distinctly from those extensively studied by Gulbransen, Phelps, and Hickman³³ in that the former were in a lower thickness range and lacked any observable crystallinity.

The electron diffraction data yielded diffuse patterns indicative of a finely crystalline state approaching an amorphous nature.⁷ The diffraction patterns showed better agreement with a hydrous metal silicate structure of the formula ($x \text{ M}_3\text{O}_4 \cdot y \text{ SiO}_2 \cdot z \text{ H}_2\text{O}$; $\text{M} = \text{Fe, Cr, or Ni}$) than a cubic spinel type although the diffuseness of the patterns does not permit a definite conclusion. The former structure is in better agreement with the compositional data. The resemblance to hydrous oxides has been previously remarked by Fontana.³⁶ It is important, however, to distinguish between hydrated oxides, hydroxides, and hydrous oxides. Water in passive oxide films does not appear to occur in stoichiometric combination nor is it lost at any critical temperature or pressure. Its presence is characteristic of water trapped during precipitation of bulk flocculent metal oxides. Moeller³⁸ and Weiser³⁹ have pointed out that such compounds are hydrous oxides differing distinctly from hydroxides and hydrated oxides. The formulation of passive film on stainless steel Type 304 can be approximated as follows:



and n is an integer, approximately 9. This formulation appears characteristic of known passive film properties and is consistent with the compositional and structural data derived from this study.

(2) *Effect of surface treatment on surface composition*

The mass of surface films is usually so small compared to that of the bulk alloy upon which they form that a selective transfer of a minute quantity of one

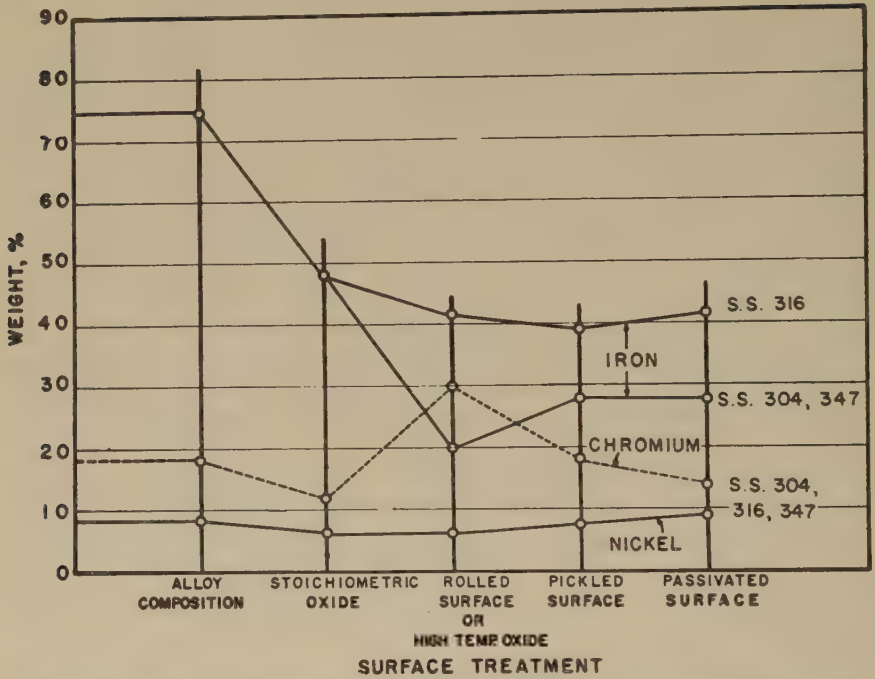


FIGURE 6. Effect of surface treatment on surface composition of major alloying elements; iron, nickel, and chromium.

or more elements to the surface can have a marked influence on film composition. Film enrichment at low temperatures may occur by selective solid diffusion, but the formation rates are more consistent with processes such as selective dissolution or selective deposition from solution. *In this way, a minor constituent of an alloy can become a major constituent of the surface.* Composition of surface films formed by various pretreatments such as cold rolling, high-temperature oxidation, pickling, and passivating were compared. Considerable redistribution between the surface and the interior was found to occur. The data for major alloying elements in film on stainless alloy Types 304, 316, and 347 are summarized in FIGURE 6. The number of samples studied and the corresponding deviations compare to those of the typical study summarized in TABLE 4. All treatments tended to reduce the iron content of the surface by approximately 40 per cent for Types 304 and 347 alloys and by 20 per cent for Type 316 alloy relative to the stoichiometric oxide. Chromium shows a marked surface enrichment after cold rolling or high-temperature oxidation. This change is apparently related to preferential diffusion of the chromium to the surface and preferential oxidation at the surface. Where this effect is unlikely, as for pickled or passivated surfaces, the chromium content of the surface is close to that of the stoichiometric oxide. The concentration of nickel in the surface remains insensitive to surface treatment. Although the surface behavior of chromium and nickel may be quite different from each

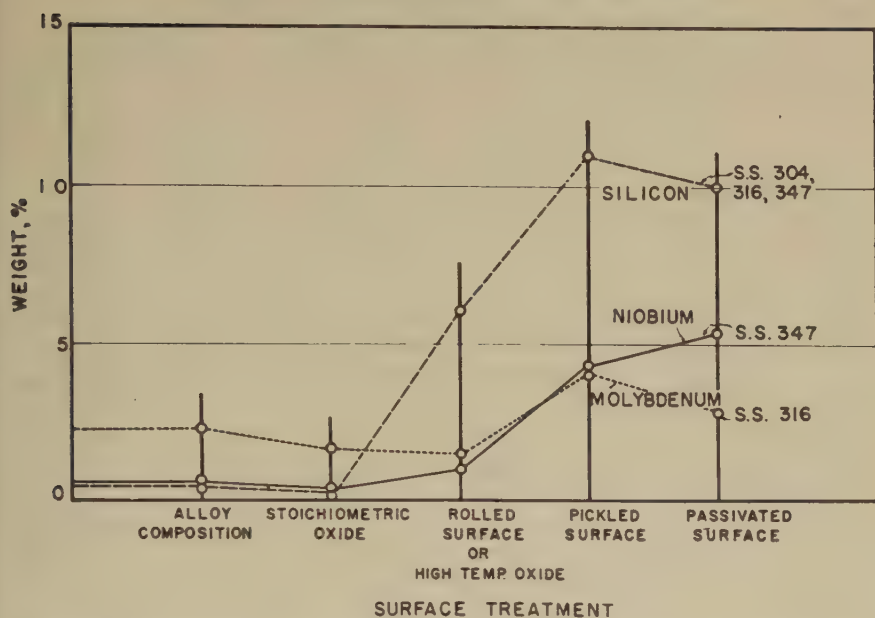


FIGURE 7. Effect of surface treatment on surface composition of minor alloying elements; silicon, niobium, and molybdenum.

other, neither metal shows any substantial enrichment effects upon pickling or passivating of stainless steel Types 304, 316, or 347. The surface composition data for three of the minor alloying elements in these three Types for the same surface treatments are summarized in FIGURE 7. The surface enrichment of silicon observed for Type 304 stainless also occurs for Types 316 and 347. It is particularly marked for the pickling operation. The resemblance of film composition data after pickling and washing to that after passivating and washing is consistent with the observation that both tend to promote pit corrosion resistance.

Niobium in Type 347 alloy also shows a definite surface enrichment upon pickling or passivation, but not on cold rolling or high-temperature oxidation. Molybdenum in Type 316 alloy does not show any substantial surface enrichment as a result of surface treatment. The conclusion may be made that the influence of chromium, nickel, and molybdenum on passivity in stainless alloy is not necessarily associated with substantial enrichment of these elements in the surface films. In contrast, passivating treatment influences substantially the quantities of iron, silicon, and niobium on the surface.

(3) Surface enrichment of silicon

The observation that silicon is enriched in passive films on stainless alloy Type 304 led to film composition studies on Type 302B stainless alloy (containing 2.5 per cent silicon) and silicon-iron alloy (containing 14.5 per cent silicon). The results, summarized in TABLE 6, compare alloy and film composi-

TABLE 6
ENRICHMENT OF SILICON IN SURFACE FILMS

Alloy	Iron	Silicon	Chromium	Nickel
	%	%	%	%
S.S. 304.....	72.0	0.6	18.3	8.7
Film.....	27.3	9.9	13.0	9.8
S.S. 302B.....	69.8	2.7	17.3	8.6
Film.....	30.0	19.0	14.3	8.0
14.5% Si-Fe.....	85.5	14.5	—	—
Film.....	37.0	35.0	—	—

tions for the three alloys. It is interesting that the iron was depleted to approximately the same value in all three films, in spite of the variations in composition.

The concentration of silicon in the surface films increased for all three alloys but is highest for the iron-silicon alloy containing the highest initial silicon. It is significant that unusually effective acid corrosion resistance of the 14.5 per cent silicon-iron alloy has been attributed to a plastic surface film composed substantially of silica.⁴⁰ The silicon enrichment in the films relative to the initial silicon in the alloy varied from 16 $\times$ for Type 304 to 3 $\times$ for the iron-silicon alloy. The data are consistent with the hypothesis that surface silica improves the acid corrosion resistance of stainless alloys, the relative enrichment in the film decreasing as the absolute silicon content of the alloy increases.

The effect of silica when incorporated homogeneously into passive films may be closely associated with the strong tendency of silica to polymerize and coagulate into stable gel structures in the presence of water solutions containing heavy metal and hydrogen ions. Hydrous silica sols are especially sensitive to coagulation by minute quantities of a great number of metal salts.⁴¹ The normal tendency of iron oxide or chromium oxide gels to be readily peptized in acid solutions is markedly reduced by the incorporation of silica into the structure. The properties of passive films are in harmony with the concept of mixed amorphous oxides as dispersions of colloid crystals or crystal fragments forming a solid gel-like mass. Berwick and Evans⁸ have demonstrated the dependence of activity-passivity limits of Type 304 stainless on the available oxygen in the system. The ability of metal oxide gels to store oxygen by adsorption of gases or ions to various degrees is consistent with this effect. In addition, the dependence of the reversibility of oxygen adsorption on surface treatments of stainless alloys is more understandable in terms of the adsorption properties of gel structures.

It is significant, however, that protective films containing silica are not always characterized by lack of observable crystallinity. Discrete particles of α -cristabolite occur on 14.5 per cent silicon-iron alloy.⁴² In addition to α -cristabolite, the presence of iron silicides has also been indicated by electron diffraction in this study of thin films isolated from 14.5 per cent silicon-iron alloys. The outstanding corrosion-resistant properties of iron-silicon alloys in hot

concentrated mineral acids may therefore depend on the degree of crystallinity of silica in the film. The validity of this hypothesis with reference to passive film properties characteristic of stainless alloys remains to be evaluated. The relationship of passivity of stainless alloys to the corrosion resistance of silicon-iron alloys may be closer than generally recognized.

Conclusions

(1) The influence of chromium, nickel, and molybdenum on passivity of stainless alloys 304, 316, and 347 is not necessarily associated with substantial enrichment of these elements in thin surface oxide films.

(2) The compositional and structural properties of thin films connected with passivity on stainless alloys can be consistently correlated with a film composed of hydrous mixed metal oxides enriched in silica.

(3) Distinctions based on compositional and structural data are made between oxide films connected with passivity and somewhat thicker oxide films formed under similar conditions.

(4) Pretreatment of stainless alloys has a marked effect on surface composition. It varies with each specific treatment but generally results in a depletion of iron and an enrichment of silicon in the surface. The influence of pretreatment on subsequent corrosion behavior is attributable to a redistribution of alloying elements between the surface and the interior.

(5) Resemblance of passive film properties to those of gel structures is observed, and a relationship to thin protective films isolated from silicon-iron alloys is suggested.

Acknowledgment

The writer wishes to acknowledge helpful discussions with E. M. Mahla, N. A. Nielsen, and with numerous members of the Engineering Research Laboratory. Much of the microanalytical work was performed by H. W. McGonigal, whose skillful assistance is gratefully acknowledged. R. H. Fisher obtained the electron metallographical and diffraction data. F. Behr, Research and Control Division, North American Philips Company, collaborated in the use of their X-ray fluorescence facilities.

References

1. UHLIG, H. H. 1948. *Corrosion Handbook*. : 22 Wiley. New York.
2. HEUMANN, T. 1951. *Z. Elektrochem.* **55**: 287.
3. EVANS, U. R. & J. STOCKDALE. 1929. *J. Chem. Soc.* : 2651.
4. VERNON, W. H. J., F. WORMWELL & T. J. NURSE. 1944. *J. Iron Steel Inst.* **150**: 81.
5. VERNON, W. H. J. 1948. Pittsburgh International Conference on Surface Reactions. : 135 Corrosion. Pittsburgh, Pa.
6. NURSE, T. J. & F. WORMWELL. 1952. *J. Applied Chem.* **2**: 550.
7. MAHLA, E. M. & N. A. NIELSEN. 1948. *Trans. Electrochem. Soc.* **93**: 1.
8. BERWICK, I. D. G. & U. R. EVANS. 1952. *J. Applied Chem.* **2**: 576.
9. MAHLA, E. M. & N. A. NIELSEN. 1946. *Trans. Electrochem. Soc.* **89**: 167.
10. MAHLA, E. M. & N. A. NIELSEN. 1948. *Trans. Electrochem. Soc.* **93**: 339.
11. SIMNAD, M. Private communication.
12. KABANOW, B., R. BURSTEIN & A. FRUMKIN. 1947. *Disc. Faraday Soc.* **1**: 259.
13. UHLIG, H. H. & S. S. LORD, JR. 1953. *J. Electrochem. Soc.* **100**: 216.
14. UHLIG, H. H. & G. E. WOODSIDE. 1953. *J. Phys. Chem.* **57**: 280.

15. WANG, H. & N. HACKERMAN. 1952. J. Phys. Chem. **56**: 771.
16. FONTANA, M. G. 1948. Trans. Electrochem. Soc. **93**: 336.
17. BATT, W. G. & H. K. ALBER. 1941. Ind. Eng. Chem. Anal. Ed. **13**: 127.
18. KOLTHOFF, I. M., T. S. LEE & D. L. LEUSSING. 1948. Anal. Chem. **20**: 985.
19. KOLTHOFF, I. M. 1948. J. Am. Chem. Soc. **70**: 2348.
20. FORTUNE, W. B. & J. G. MELLON. 1938. Ind. Eng. Chem. Anal. Ed. **10**: 60.
21. ALEXANDER, O. R., E. M. GODAR, & N. J. LINDE. 1946. Ind. Eng. Chem. Anal. Ed. **18**: 206.
22. VAN DER WALT, C. F. J. & A. J. VAN DER MERWE. 1938. Analyst. **63**: 809.
23. URONE, P. F. & H. K. ANDERS. 1950. Anal. Chem. **22**: 1317.
24. WILLARD, H. H. & L. H. GREATHOUSE. 1917. J. Am. Chem. Soc. **39**: 2366.
25. RICHARDS, M. B. 1930. Analyst. **55**: 554.
26. BOLTZ, D. F. & M. G. MELLON. 1947. Ind. Eng. Chem. Anal. Ed. **19**: 873.
27. BUNTING, W. E. 1944. Ind. Eng. Chem. Anal. Ed. **16**: 612.
28. DEVEREAUX, H. D. & G. E. MOORE. To be published.
29. LAUW, A. B. H., S. S. LORD, JR., & D. N. HUME. 1952. Anal. Chem. **23**: 1169.
30. FREUND, H. & A. E. LEVITT. 1951. Anal. Chem. **23**: 1813.
31. SANDELL, E. B. 1936. Ind. Eng. Chem. Anal. Ed. **8**: 336.
32. HURD, L. C. & H. O. ALLEN. 1935. Ind. Eng. Chem. Anal. Ed. **7**: 396.
33. GULBRANSEN, E. A., R. T. PHELPS, & J. W. HICKMAN. 1940. Ind. Eng. Chem. **18**: 391; 1946. *Ibid.* **18**: 640.
34. TOKUMITU, T. 1940. Nature. **145**: 489.
35. MIYAKE, N. 1937. Rep. Inst. Phys. Chem. Res. Japan. **31**: 161.
36. FONTANA, M. G. 1947. Corrosion. **3**: 567.
37. ALEXANDER, J. 1937. Colloid Chemistry. 4th ed. : 259. Van Nostrand. New York.
38. MOELLER, J. 1953. Inorganic Chemistry. Chap. 4. Wiley. New York.
39. WEISER, H. B. 1935. Inorganic Chemistry. II. (1) Wiley. New York.
40. MCKAY, R. J. & R. WORTHINGTON. 1937. Corrosion Resistance of Metals and Alloys. : 277. Reinhold. New York.
41. ALEXANDER, J. 1937. Colloid Chemistry. 4th ed. : 144. Van Nostrand. New York.
42. YAMAGUCHI, S. 1941. Bull. Chem. Soc. Japan. **16**: 332.

SURFACE STUDIES WITH THE ELECTRON MICROSCOPE

By C. J. Calbick

Bell Telephone Laboratories, Murray Hill, N. J.

Introduction

In the field-emission microscope discussed in a previous paper in this monograph by Doctor Becker, the topological and other properties of surfaces are deduced from the electron emission patterns observed. Emission microscopes using thermionic, photoelectric, or secondary electrons also have been used in a similar manner, although with magnification greatly inferior to that of the field-emission microscope. Valuable information on a few types of surfaces has been obtained, but most surfaces are not suitable for emission studies. The more common electron microscope is a transmission-type instrument similar to the biological optical microscope. The electron microscopist finds, however, that practically all materials which he examines are opaque to electrons and can be seen only in profile. To observe a surface, he resorts to representing it by a thin membrane of locally varying thickness called a replica. The simplest example is a collodion membrane formed by allowing a dilute solution to dry on the surface. This membrane will be thick in the valleys and thin on the hills of the surface. When it is removed from the surface, supported on a fine-mesh screen and examined in the microscope, the thicker portions scatter more electrons out of the primary beam. *Absorption* is negligible for the usual thicknesses of replica films. In fact, single scattering without loss of energy predominates, although some plural scattering may occur in the thicker parts of the replica. If the lens were perfect and there were no objective aperture, all the electrons, scattered and unscattered alike, would be focused on the image plane, which would be illuminated to uniform brightness. The actual lens suffers from a high degree of spherical aberration. In consequence, the scattered electrons are not focused on the image plane, but form a diffuse background.^{1, 2} This background can be largely eliminated by the use of an objective aperture. The pictures, accordingly, are images of the primary radiation from which a varying amount has been locally removed.

Resolution: Instrumental and Replica

According to the Abbé theory, the instrumental resolution is proportional to the wave length and inversely proportional to the numerical aperture. The latter is determined by the angular divergence of the radiation incident upon the objective lens. In the optical microscope, the angular divergence is normally fixed by the aperture of the objective lens itself (at least in the case of "critical illumination"). In the electron microscope, the angular divergence is usually determined by the illuminating system, *i.e.*, by the cathode and its associated lenses, and by the condenser lens. A very small value, about 0.003 radians, is required on account of the large spherical aberration of the objective lens. Now λ is about 0.05 Å. (It ranges from 0.07 Å for 30 kv. to 0.039 Å for 100 kv. electrons.) The instrumental resolution by the Abbé theory there-

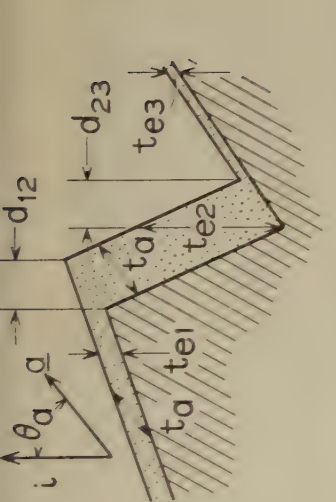
fore is about $10 \text{ \AA}$. W. Glaser³ has given a wave-mechanical treatment based upon a bell-shaped field which indicates a minimum value $d \simeq 4 \text{ \AA}$. In this analysis, the axial magnetic field is represented by an inverse quadratic function, for which the electron-optical equations are solvable. In practice Hillier⁴ has been able to realize a resolution of about $10 \text{ \AA}$. This extreme resolution is possible only if great care is exercised to provide axially symmetric (anastigmatic) electron lenses. Moreover, it is difficult to provide objects suitable for testing resolution which is approaching the size of individual atoms. In the ordinary practice of electron microscopy, the features of interest are usually on a considerably coarser scale.

The very small numerical aperture, about $10'$ of arc, results in great depth of focus, about 20μ . Very rough surfaces can therefore be studied. Also, the great depth of focus permits tilting the object to two different angles. This potentiality yields stereoscopic pairs of micrographs with the fundamental advantage of three-dimensional visualization, even at the highest magnifications.

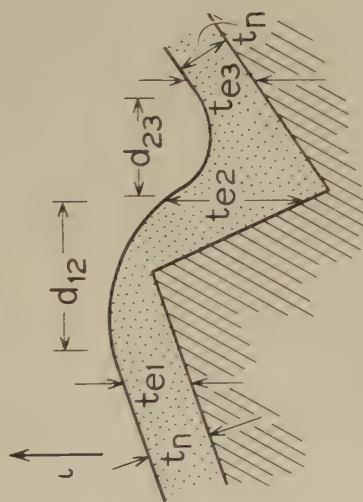
The representation of a surface by a replica results in a loss of resolution, since a sharp corner is not reproduced as an equally sharp change in film thickness. FIGURE 1 illustrates the various possibilities of replica geometry at such corners. Ideally, a sharp corner would be represented by a replica of the geometry shown in FIGURE 1 (a), with both surfaces of the replica exhibiting sharp corners. It is barely possible* that surface oxidation, for example as of aluminum, produces an oxide skin of this kind. Even in this case, however, the thickness t_e in the direction i of the electron beam changes from a value characteristic of one slope to that characteristic of another over a finite horizontal distance d . If material evaporating from a point source condenses with every atom sticking where it strikes the surface, the film geometry of FIGURE 1 (b) results. Most metals condense in this fashion. It can be readily shown that d_{12} is proportional to $t_{e2} - t_{e1}$. Thus a local difference in replica thickness always entails a finite *intrinsic resolution* d of the replica film.⁵ If the evaporated material wanders over the surface as a two-dimensional gas before condensing, the film geometry of FIGURE 1 (c) is believed to result. Here the local thickness t_n is constant except for rounding at the corners. Silica and several other oxides form films of this type in shadowed areas, that is, areas shielded from the source by intervening obstacles. The fact that sharp shadow edges are also formed (FIGURES 4 to 11) indicates that part of the condensing material sticks where it strikes. So silica and other evaporated oxides form films with geometries intermediate between FIGURES 1 (b) and 1 (c). Finally, FIGURE 1 (d) is diagrammatic for a plastic replica. The curvature of the upper surface has been simply sketched in, since there is no known basis for computing it.

The intrinsic resolution of the replica varies locally, being dependent on surface topography in relation to the electron beam and, in case (b), also to the direction of the atom source. It is proportional to the average linear thickness of the replica, and hence inversely to the density of the replicating material.

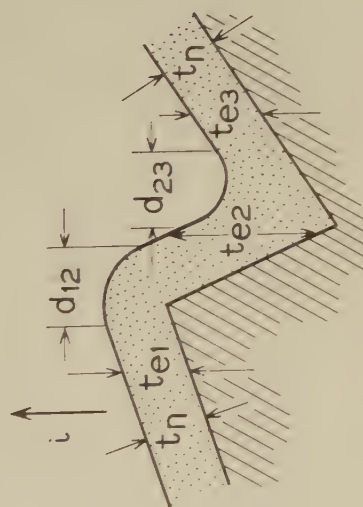
* More probably, its geometry is that of FIGURE 1 (c), with the shading lines indicating base material located above rather than below the replica, shown as the dotted strip.



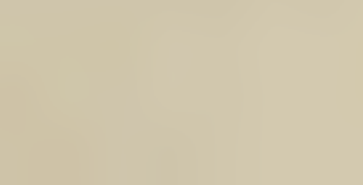
(a) t_n CONSTANT, NO ROUNDING



(b) t_a CONSTANT: ATOM SOURCE IN DIRECTION $\underline{a}$



(c) t_n CONSTANT, CORNERS ROUNDED



(d) PLASTIC FILM: DIAGRAMMATIC ONLY
 FIGURE 1. Intrinsic resolution in direction i . Replicas represent sharp corners by changes in thickness spread over distances d which depend on slope relative to i ; in (b) d also depends on α the direction of the atom source.

TABLE 1
RESOLUTIONS OF REPLICA FILMS, 10 $\mu\text{g}/\text{cm}^2$

Material	Density	$t_e, \text{\AA}$	Resolution, $\text{\AA}$
Plastics*	1.05	1000 (approx.)	500 (approx.)
Si, SiO_2 , SiO , Al.....	2.4, 2.5, 2.5, 2.7	420-370	210-185
Al_2O_3 , TiO_2 , GeO_2 , Ge.....	3.9, 4.0, 5.0, 5.4	260-180	130-90
Cr, NiO , Ni.....	6.9, 7.5, 8.9	145-112	73-61
Ag, U_2O_3	10.5, 10.5	95	48
Gold-manganin.....	10-17	100-60	50-30
U_3 , Au, Pt.....	18.7, 19.3, 21.7	54-46	27-23

* For comparison only, resolution not calculable.

† On exposure to air, U oxidizes.

The average thickness determines the range of tone values which can occur in the micrograph, and may be compared with the exposure of a photographic negative. A thin replica has high resolving power, but can produce only a limited range of contrast. The idea of an optimum average thickness, analogous to optimum exposure of a photographic negative, arises naturally. It depends on the surface being portrayed, but is often in the range 5 to 10 $\mu\text{g}/\text{cm}^2$.

Although the resolution varies locally, it almost always lies in the range from $\frac{1}{4}$ to 2 times the average thickness. It is reasonable to ascribe to d a value $\frac{1}{2}$ the average thickness, for purposes of comparison. Resolutions of replicas of optimum thickness thus range from about 500 $\text{\AA}$ for plastic replicas, down to 25 $\text{\AA}$ for the heavy metals (TABLE 1). In practice, instrumental resolution usually lies in the range of 20 to 50 $\text{\AA}$, the possible resolution of 10 $\text{\AA}$ being degraded by slight astigmatic and other aberrations. The effective resolution may be taken as the sum of the instrumental and the intrinsic resolutions. Up to the present, it has rarely been less than 100 $\text{\AA}$, and often considerably greater.

Moreover, many replicas are made in a two-step process in which a plastic mold of the surface is made first. If the plastic, as shown in FIGURE 1 (d), is quite thin, of order 100 $\text{\AA}$ in thickness, and metal or other inorganic material is subsequently evaporated onto the plastic surface after removal from the original surface, we have a "shadowed plastic replica," which is much used. Again, the plastic may be not thin, but very thick, as for example by pressure-temperature molding into a block of plastic. The specimen is subsequently separated mechanically from this block. An inorganic film is then deposited by evaporation upon the plastic, from which it is later freed by dissolution of the plastic. When an intermediate plastic replica is used, a granular structure on a scale of the order of 100 $\text{\AA}$, attributable to the large plastic molecules, is found if the final replica has sufficient intrinsic resolving power. The granular structure is shown in FIGURE 2. It can be avoided by evaporating the inorganic film directly onto the original surface. Serious difficulty may be encountered in removing this film from the surface. If a thin plastic film is used for this purpose, the replica is called a preshadowed plastic replica. Since the plastic molecules themselves are unshadowed, their contrast is insufficient to produce plastic granularity in the image.



FIGURE 2. Germanium replica, $\theta_n = 35^\circ$, of sintered NiMn oxide. Resolution is better than $80 \text{ \AA}$ width of micron line. Molecular structure of intermediate plastic mold is evident. Thin film in shadow.

Shading and Shadowing

When the atom source during the evaporation is in a direction considerably different from normal as in FIGURE 1 (b), the thickness variations in the film result in density variations in the micrograph which may be termed shading. Prominent features may prevent the exposure of certain parts of the surface directly to the source. These parts appear as shadows in the micrograph. Within the shadows, shading or replication does not occur unless there is a surface-mobile component as in silica and other oxides. Detailed interpretation of shadows is, in general, not possible, since their sizes and shapes depend both on the profile casting the shadow, and on the topography of the terrain where the shadow falls. In cases where the profile has a regular geometry and the terrain is flat, shadows give very useful measurements of height. The shadow profile may also furnish one boundary useful for the measurement of horizontal extent of the feature. On many surfaces, the number and extent of shadows increase rapidly as the polar angle θ_n [FIGURE 1 (b)] increases.

The interpretation of electron micrographs is dependent in various degrees on all the physical factors rather inadequately discussed above. A great simplification can be obtained by using stereoscopic pairs. The three-dimensional view very often yields immediately the information desired. Moreover, it may demonstrate clearly some error in the two-dimensional interpretation of single micrographs.

Surfaces of Particles

In the study of particles, the usual practice is to disperse them in a plastic solution from which a thin plastic film is made. They may also be dispersed on top of a plastic film, by one or more of several methods. These include simple dusting, spraying from a liquid suspension, electrostatic and supersonic methods, and several others. Filmless mounting, in which the particles adhere to mesh screen or other web-like supports is sometimes used. Since the particles are usually opaque to the electron beam only profiles can be observed.

Particles also can be represented by surface replicas. The surface-migratory behavior of condensing silica makes it especially suitable for this purpose. In this technique, the particles are dispersed on a plastic film on a microscope slide. The slide is mounted on the silica evaporating unit shown in FIGURE 3 and a layer of silica deposited by evaporation *in vacuo*. The slide is then turned end for end and a second evaporation carried out. In the unit shown, the polar angle of incidence is 60° ; the azimuths of the two evaporations differ, of course, by 180° . After removal from the evaporator, the plastic film is stripped from the slide and supported by a suitable mesh screen (if the plastic film were not used, the silica film could not be stripped from the slide). The plastic is then dissolved,* leaving the silica film with its embedded particles adhering to the screen. The left-hand picture of FIGURE 4 shows a micrograph of some alkaline-earth oxide particles. The particles are here represented by profiles and attendant shadows. The next step is to dissolve* the particles themselves, leaving only the silica shells to represent the particles, as seen in the right-hand picture.

The question is often asked, how badly are replica films distorted by surface tension forces during the processing? In the case of FIGURE 4, the two pictures actually constitute a stereoscopic pair. Quite evidently, the replica has not been greatly distorted by the intervening wetting with dilute acid, and by subsequent drying. But a word of caution is necessary. Replica damage can occur locally. In the right-hand picture of FIGURE 4, two extra particle shells are visible (arrows). One of these particle shells had torn loose at its base, leaving a hole in the film (with which it was identified) in a nearby region not in the field of the picture. It had drifted about 20μ across the replica surface. Only a small percentage of particle shells tear loose in this fashion. Stereoscopic observation, best done on the original plates, indicates that they do not collapse. This is attributable, in part at least, to the stiffening effect of three-dimensional fine surface structure. With highly prolate smooth-surface par-

* Least damage is done to the silica film if dissolution is carried out at the solvent surface.

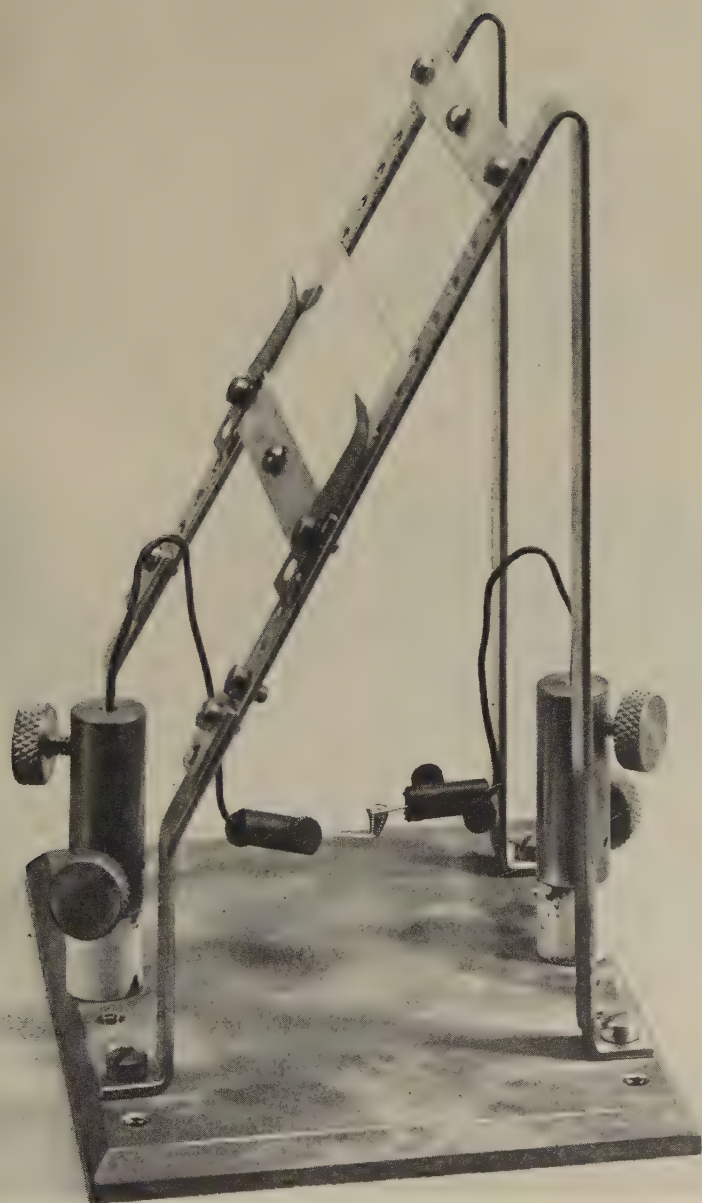


FIGURE 3. Unit for evaporating silica or other similar materials.

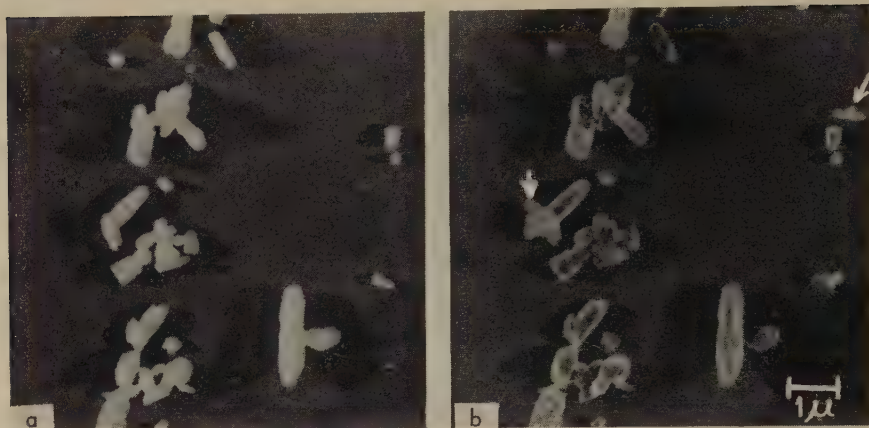


FIGURE 4. (Ba Sr Ca) CO_3 particles. (a) Supported by and encased in silica; (b) same field, silica shells after dissolving particles. Arrows show loose silica shells from adjacent regions.

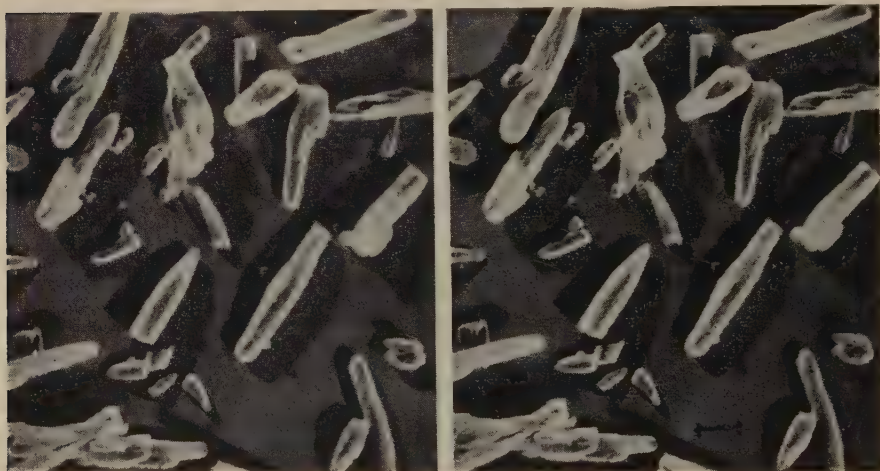


FIGURE 5. Fragmentation of prolate particles. Flat ends are thought to indicate broken particles.

ticles and, in particular, with fibers larger than 1 or 2μ in diameter, replica collapse can become a serious problem.

FIGURE 5 shows a stereoscopic pair in which the flat ends of several particles are believed to indicate fragmentation. One particle shows an offset, thought to be caused by transverse fracture. It was found almost impossible to disperse these prolate particles without fragmentation. Less prolate particles of similar materials, such as those of FIGURE 4, seldom show flat ends. Silica shell replicas also permit ascertainment of the degree of aggregation, and the identification of particles in aggregates when the latter are not more than two or

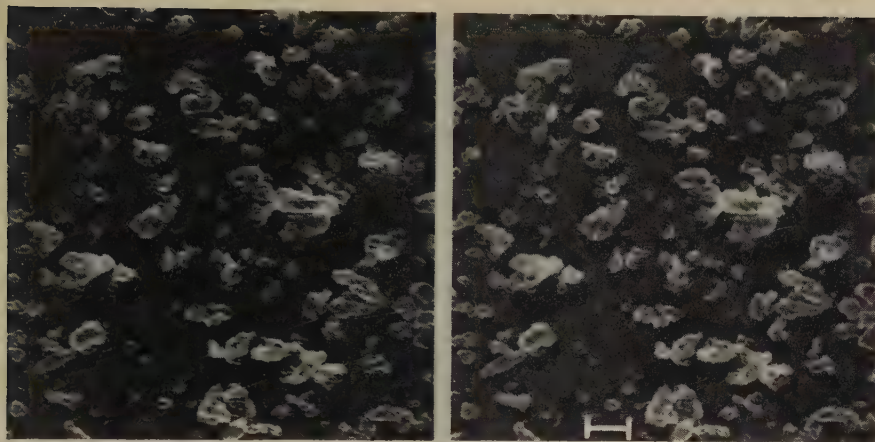


FIGURE 6. Smaller particles have not been well dispersed, but are usually identifiable in aggregates.

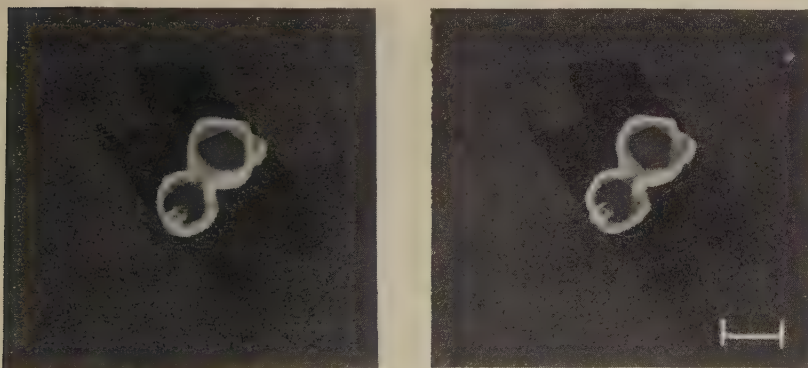


FIGURE 7. Two octahedral units of carbonyl nickel physically joined. Beginning of reticulate growth character.

three particles thick. The particles of FIGURE 6 are smaller and hence more difficult to disperse than those of FIGURE 5. They are considerably less prolate, in consequence of which fragmentation is not observed. But aggregation is rather general. Particles in aggregates can be distinguished and measured. Of course, one cannot be *sure* that these particles —or some of them, at least —are not physically joined together and constitute therefore a single particle having a reticulate structure. Cases of branching-type particles were sometimes observed in the powder illustrated in FIGURE 5. Careful study of the original stereomicrographs* usually permits a reasonable determination of whether the particles are physically joined together, or are merely adhering together in an aggregate.

The nickel powder illustrated in FIGURES 7, 8, and 9 is illustrative of physical

* Unfortunately, the halftone method of reproduction tends to destroy the stereoscopic effect.

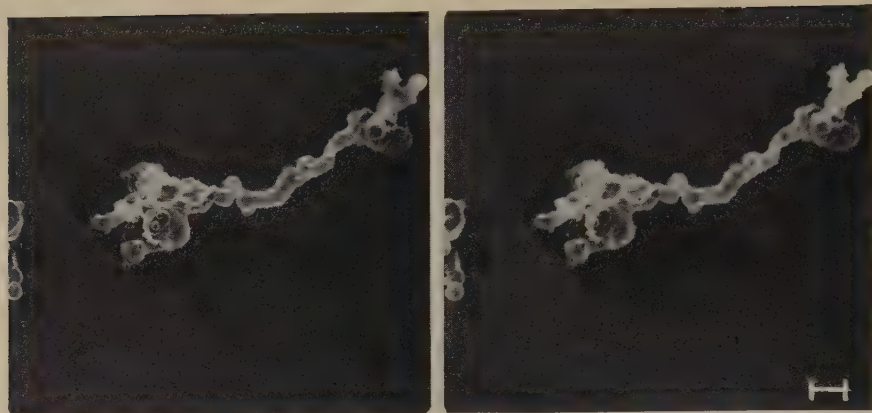


FIGURE 8. Carbonyl nickel "particle." Reticulate growth produces erratic shape.

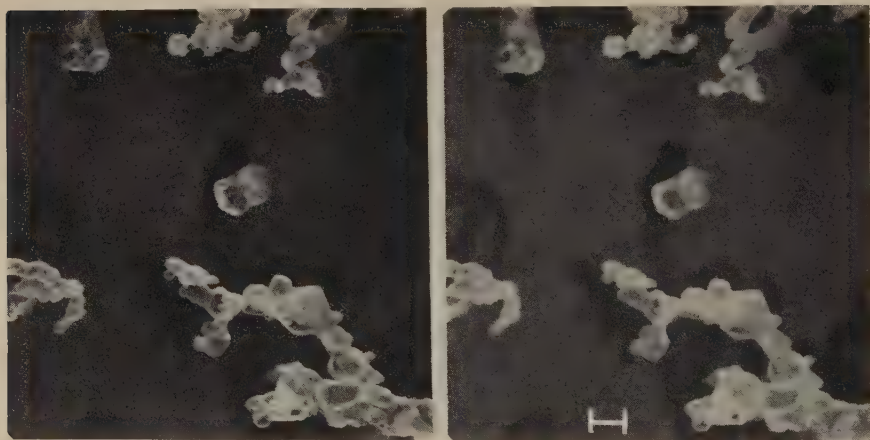


FIGURE 9. Stereoscopic pair show several cases of tetrahedral and octahedral surface development.

joining together of particles. This powder and that illustrated in FIGURES 10 and 11 were grown by the carbonyl process. The surfaces in general tend to be the tetrahedral or octahedral (111) planes. FIGURES 7, 8, and 9 all show a characteristic reticulated growth of particles of the first powder. This reticulation is most clearly seen in FIGURE 7, where two nearly octahedral particles are growing together at a small angle, as a sort of bi-crystal. Extension of this process results in the growth of the long reticulated particles (FIGURES 8, 9). In some cases, these particles terminate in tetrahedral units. Aggregation has frequently been observed, but is not illustrated here. This powder had particle size of 1.5μ determined by the nitrogen adsorption method. The micrographs verify the surface area, since the individual blocks or units of the reticulated structure are of order 1μ in size.

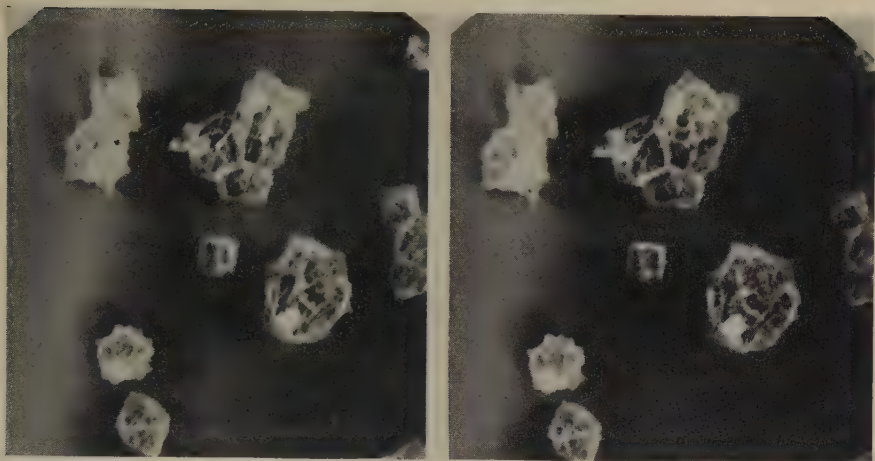


FIGURE 10. A different carbonyl nickel powder. Particles of either cubiform or rounded shape aggregate but are not usually physically joined together.

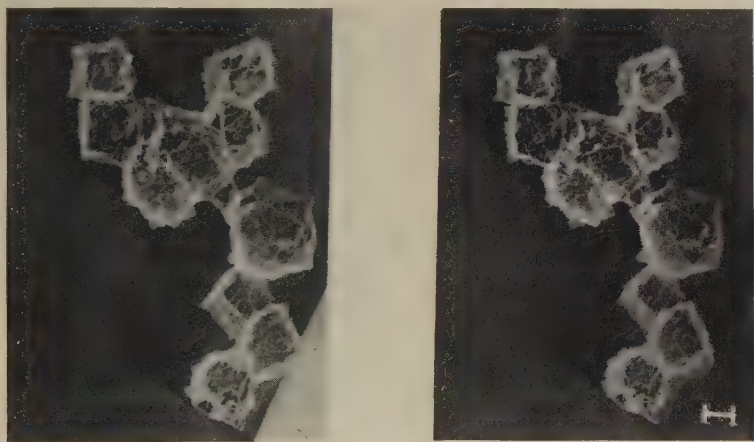


FIGURE 11. Stereoscopic observation on cubiform particles shows that spires which are distributed over the surface tend to have (111) surfaces.

The second nickel powder (FIGURES 10 and 11) showed none of this reticulated nature. The particles were either cubiform or rounded. FIGURE 10 shows particles of each type. They were covered with spires, the surfaces of which one might assume to be the tetrahedral planes of the crystal. Stereoscopically, it can be often observed that the surfaces of the spires on a single particle are approximately parallel in groups. The numbers of these groups would be four, if (111) planes alone were developed. Since the spires have somewhat rounded sides and corners, and their sides are small, stereoscopic determination of angles suffers from considerable uncertainty. So, while it is believed to be highly probable that the spires are bounded by approximately (111) planes, it may be that other planes also occur.

It is obvious that the percentage of voids in any compacted forms of these powders will be rather large. The powders will also have other bulk properties considerably different from those of the equivalent spheres conventionally discussed in powder technology.

The limitation that the silica shell must not be attacked during dissolution of the particles prevents the examination of many powders by this method, particularly of those used in ceramic technology. Although surface replica shells can be made of many other materials, the most satisfactory are probably the oxides such as TiO_2 , GeO_2 , *etc.*, which condense with surface diffusing components. None of these is superior to silica in resistance to chemical attack.

Germanium Surfaces

Etch surfaces of germanium single crystals rarely display much structure of interest, beyond a few etch pits. However, along a boundary between two parts of the crystal differing in orientation by a minute of arc or less, Vogel, Pfann, Corey, and Thomas⁶ have recently observed regularly spaced depressions. The difference in orientation is much less than can be detected as lineage in ordinary X-ray measurements. Special X-ray techniques were used to detect and measure the slight displacements of the reflected spot owing, in the reported cases, to orientation differences of $17.5''$, $65''$, and $85''$, respectively. Theory indicates that along the boundary dislocations should be present and spaced at a separation $s = d/\Delta\theta$, where d is the distance between atoms in the direction of the misorientation boundary, and $\Delta\theta$ is the orientation difference in radians. The line of dislocations was along the trace of a (110) plane, so $d = 4 \text{ \AA}$. The three cases reported gave calculated values of $s = 4.7$, 1.3 , and 0.97μ . The observed values were 5.3 , 1.3 , and 0.99μ , agreeing with the calculated values well within the experimental uncertainty.

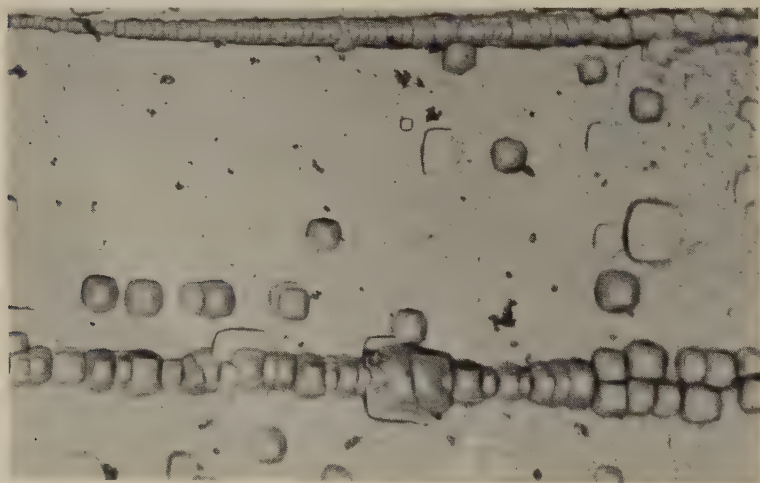


FIGURE 12. Etch pits in (100) germanium surface owing to dislocations along misorientation boundaries. Optical micrograph.



FIGURE 13. Electron micrograph shows details of a few pits along upper row in FIGURE 12. Calculated misorientation is 14 seconds.

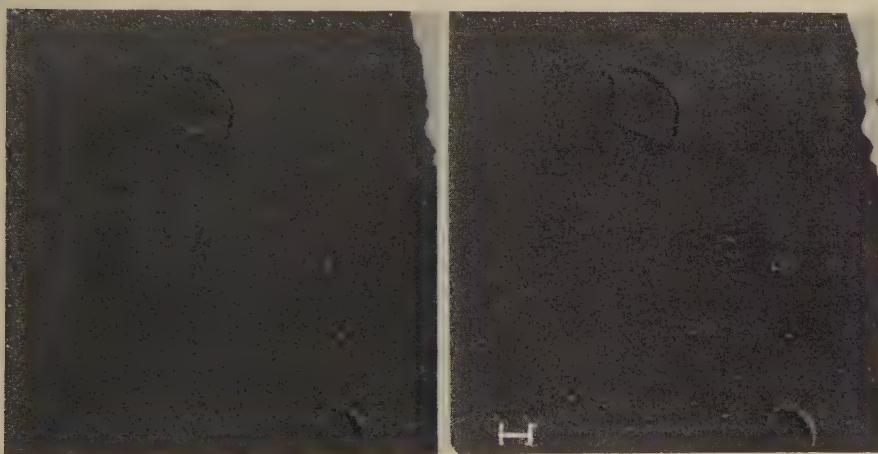


FIGURE 14. Details of pits along lower row in FIGURE 12. Calculated misorientation is 7 seconds.

FIGURE 12 is an optical micrograph of an area in a different crystal showing two lines of dislocation markings, the upper one characterized by a spacing $s \simeq 6\mu$, the lower by $s \simeq 12\mu$. Correspondingly, $\Delta\theta = 14''$ and $7''$. There appear also sporadic other markings of a similar nature, and a few etch pits. This surface was rather heavily etched, to bring out the sharp nature of the outer boundary, which is a rather surprising and as yet unexplained feature. FIGURES 13 and 14 are electron micrographs of markings found in the same region, but with a somewhat lighter etch. The central cores are seen to be narrow wing-shaped depressions (FIGURE 13), or quadrilateral (FIGURE 14). The upper line of the optical micrograph (FIGURE 12) shows V-shaped cores, while at least some of the cores of the lower line show a distorted fourfold symmetry. The very small misorientation ($7''$) thus appears only to distort the fourfold symmetry of the (100) face. The larger misorientation ($14''$) reduces the symmetry to roughly twofold in the boundary region. One naturally ascribes the markings to changes in the electrochemical potential in the vicinity

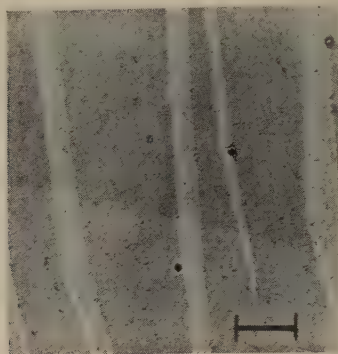


FIGURE 15. Smooth area of a cleavage surface of germanium.

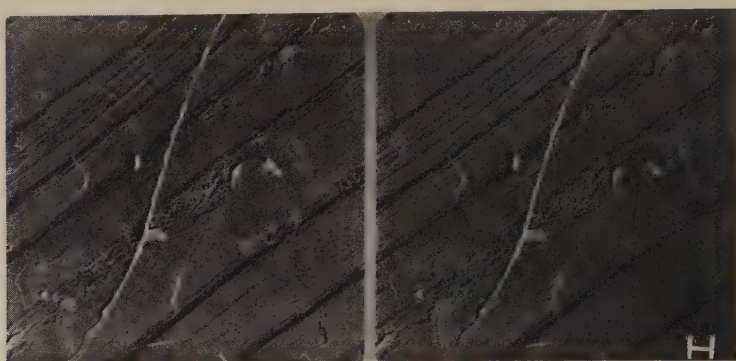


FIGURE 16. Electron micrograph showing crack crossing a series of cleavage steps. Some inclusions are evident.

of the dislocation. The distribution of these changes as indicated by the shapes of the markings is more difficult to explain. Lines of markings of similar origin are found also on (111) surfaces.

Cleavage surfaces of germanium crystals also show interesting structure. Much of the area is quite smooth (FIGURE 15). Even these areas show structure in the form of fracture striae, which appear in both optical and electron micrographs. Cleavage steps appear, and FIGURE 16 shows a case where a surface crack traverses a number of such steps. Dependent on the particular crystal examined 75 to 95 per cent of the area has these characteristics. The remainder shows more numerous cracks. FIGURE 17 portrays such an area.

In the process of replication, the two-step process using an intermediate plastic mold was used. A crack in the germanium therefore appears as a fin on the plastic mold. Furthermore, if the crack is quite narrow compared with its depth, the depth of penetration during molding is determined by its width, the viscosity of the plastic, and the pressure applied during molding. Variable width of a crack would result in variable depth of penetration. The scalloped edges of some of the fins are believed to be owing to this effect and to indicate variable width rather than variable depth of crack. Chromium was evaporated

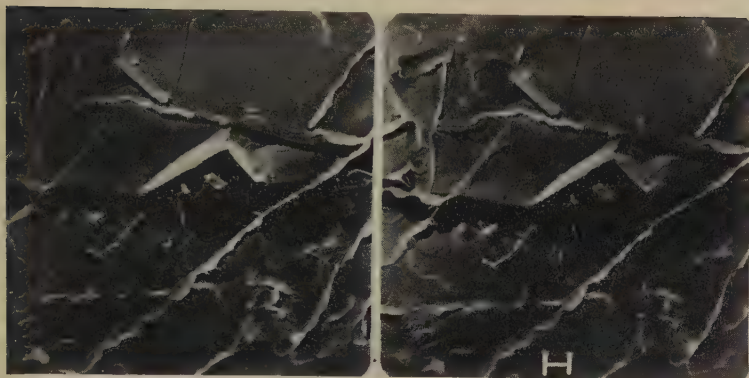


FIGURE 17. Electron micrograph showing a region with more numerous cracks.

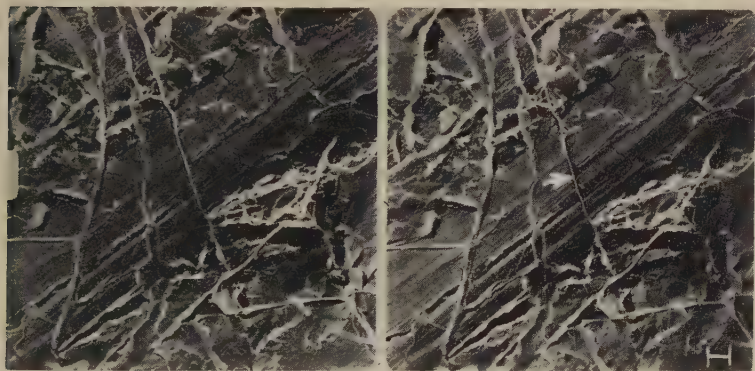


FIGURE 18. One crack (arrow) was almost in the plane of incidence of the source during replica formation. Its width is about 200 Å.

onto the plastic mold surface at polar angle of incidence of 35° . This chromium film was the replica from which the micrographs were obtained.

One of the fins (cracks) in FIGURE 18 is almost parallel to the azimuth of incidence of the chromium. The width of this fin indicates a value of about 200 Å for the width of the crack. There is considerable uncertainty in this measurement, for the chromium itself is more than 100 Å thick, and about all we can say is that the width lies between 100 Å and 300 Å. This, however, is just the range where variable width would result in variable penetration for the large plastic molecules. The fact that they penetrate it at all to form the fin indicates that the crack is more than about 50 Å wide. It is almost certain that smaller cracks would not be replicated.

FIGURE 19 shows several pairs of parallel cracks, and also single cracks (not paralleled by others within the field of view). The angles between the crack directions can be correlated with those between traces of low-index planes. In particular, if the vertical crack is regarded as the trace of a (110) plane, most of the other cracks make angles of approximately 11° , 14° , 19° , and 30° with



FIGURE 19. A region showing a few parallel pairs of cracks. Vertical crack is probably the trace of a (110) plane and other cracks are approximately traces of planes of low index.

this trace or the two other (110) traces on the (111) surface. These are the angles between traces of low-index planes. Many years ago, Andrade and Martindale,⁷ experimenting with thin films of silver and gold sputtered onto diamond cleavage surfaces, observed that these films coagulated into spherulets, on heating to 260° to 280° C. for silver and 380° to 400° C. for gold. These spherulets formed lines, the angles between which corresponded to those between traces of low-index planes on the (111) cleavage surface. They attributed these lines to cracks in the surface. Similar lines, but of random orientation, were found on surfaces of fused quartz. The character of the coagulation was quite different from that associated with even the lightest scratch artificially applied to the quartz. Andrade and Martindale observed the surface under glancing illumination in strong light. The spherulets scattered the light, and were intensely bright. They calculated that a crack connecting the linearly-disposed spherulets should scatter enough light to be visible, if it were as much as $50 \text{ \AA}$ wide. Lines connecting spherulets were not observed. Andrade and Martindale adduced considerable auxiliary evidence as to the reality of the cracks for which the reader should refer to the original paper. They concluded that the cracks exist but are less than 10 atoms across, and they suggest that these fissures are Griffith cracks. Such narrow cracks would not be observable with the present electron-microscopic techniques. We have examined diamond surfaces, but have seen no features that could be certainly identified as cracks, although some faint lines *possibly* of this origin were observed.

In the case of germanium, a possible explanation of the genesis of the cracks follows. Seitz⁸ has proposed, as an explanation for the very low-temperature annealing of resistance changes induced by cold-work in certain metals, that the vacancies created by cold-work (of order 10^{18} per cm^3) have sufficient low-temperature mobility to coagulate into sheets of vacancies, probably nucleated at imperfections. These sheets scatter electrons less than the same vacancies randomly distributed in the lattice, and hence the coagulation process would reduce the resistivity. This resistance annealing proceeds rapidly at room temperature in the case of copper, and at a measurable rate as low as -80° C.

Now our germanium crystal is formed by drawing from the melt and is cooled during the process by jets of hydrogen. It exhibits some ductility down to perhaps $600^{\circ}\text{C}.$; below this temperature it becomes brittle. Thermally induced vacancies could coagulate into sheets. During the cataclysmic process of cleavage, those sheets crossing the fracture surface would enlarge to relieve strains, if stress were present because of uneven cooling. To explain the results, it is then only necessary to assume that the bulk of the crystal is sufficiently strain-free so that only fracture striae are formed. Dependent on the previous history of the crystal, there may be regions in a condition of strain which can be relieved by the enlargement of the vacancy sheets to form the observed cracks.

The fracture strength of germanium is very low. Frequently, when germanium and plastic are molded together in the process of making replicas, and are then mechanically separated, fracture occurs within the germanium rather than at the germanium-plastic interface. It is well known that all materials exhibit fracture strengths orders of magnitude less than the theoretical values. In 1920, Griffith⁹ proposed an explanation based on the assumption that all materials have internal defects consisting of microcracks. At the peripheries of these cracks, the stress would exceed the average stress by amounts depending on the size of the crack transverse to the direction of tensile stress. Fracture would then occur when the theoretical strength of the material was exceeded at the boundary of the largest or critical crack. By rapidly drawing glass and quartz fibers, the fracture strength was greatly increased. This increase was attributed to the rapid elongation and consequent narrowing down of the voids transversely to the fiber axis. Physicists and metallurgists have as yet found no adequate substitute for the Griffith theory of brittle fracture.^{10, 11} The vacancy sheets of Seitz and the Griffith microcracks are certainly very similar physically.

Recently Hoffman, Anders, and Crittenden¹² have adduced some further experimental evidence in support of Seitz' ideas. Nickel films evaporated onto $0.002''$ strips of mica are in a state of high initial tension which is measured by the curvature it causes in the mica strip. Annealing for periods of 30' at successively higher temperature results in a decrease in the tensile stress up to about $200^{\circ}\text{C}.$, and thereafter in an increase. The authors attribute this behavior to formation of vacancy sheets with an increase in volume, followed by collapse of these sheets with decrease of volume at the higher temperatures.

If Griffith microcracks which, according to Roop¹⁰ "exist in great numbers and various orientations in every specimen of metal" likewise exist in germanium, and are the sites for the larger cracks observed, their volume density must be at least $10^{12}/\text{cm}^3$. The maximum density of observed cracks is of order $10^8/\text{cm}^2$, which corresponds to a volume density of $10^{12}/\text{cm}^3$. Such a density of vacancy sheets would be expected to affect the electrical properties in a measurable manner, but no such effect has been detected. This lack of effect casts serious doubt on their existence. Also, cracks are not found on etched surfaces of germanium. Since etching would not cause cataclysmic enlargement of the minute cracks, one would not expect to observe them directly. But if single dislocations can affect the electro-chemical potential

and produce the pits seen in FIGURES 12, 13, and 14 vacancy sheets should likewise have observable—and presumably much greater—effects (this argument is not conclusive, because the greater effects could so overlap each other as to average out). Since impurity centers or other crystal imperfections could presumably also act as sites for the initiation of cracking during cleavage, the observed cracks are by no means conclusive evidence for the existence of Griffith's microcracks or Seitz's vacancy sheets.

Pores in a Substitutional Metal Alloy

A nickel-iron permalloy was electroplated with about 10μ of gold, and subsequently heat-treated at 800°C . to stabilize its magnetic properties. Before heat treatment, the plated surface had a pebbly texture, with occasional markings presumably owing to grains of the base metal. Subsequent to heating, the surface was characterized by distributed pores in a great variety of shapes (FIGURES 20 and 21). Examined by electron diffraction, the outer surface yielded a face-centered pattern with a constant in the range of the Ni-Fe alloys (3.52 to $3.57\text{ \AA}$). Treatment with 10 per cent HCl removed this layer, and a face-centered structure in the range of the Au-Fe and Au-Ni alloys was found. X-rays yielded a more precise value, $3.99\text{ \AA}$. This value compares with $4.07\text{ \AA}$ for gold. These experimental facts indicate that (1) during heat treatment nickel and probably also iron dissolved in the gold; (2) during cooling a very thin layer of nickel, a few millimicrons thick, precipitated on the outer surface; (3) vacancies coagulated to form pores in the alloy layer.

The vacancies may have been present in the gold layer *ab initio*. Electroplated layers are usually in a state of strain, and coagulation of vacancies could occur during the heat treatment. Or the vacancies could originate in the manner discussed by Baluffi and Alexander,¹³ who observed porosity in diffusion couples of copper- α -brass, copper-nickel and gold-silver. Baluffi and Alexander conclude that porosity may develop in any substitutional metal

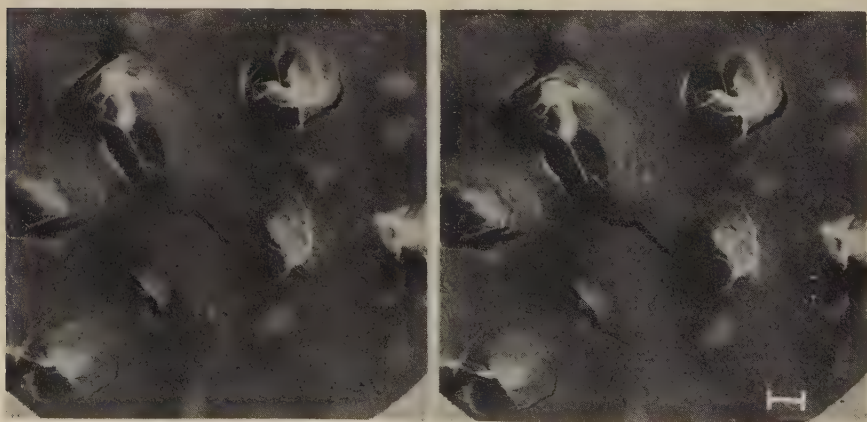


FIGURE 20. Pores developed by heat treatment of gold-plated permalloy.

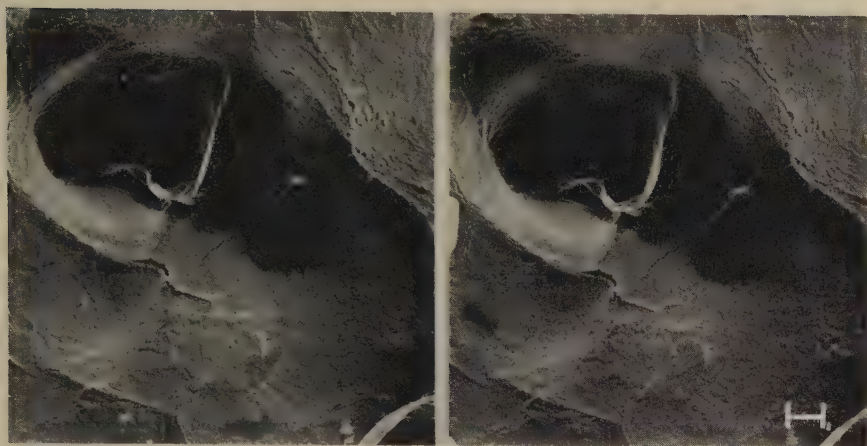


FIGURE 21. A very large pore. Depth here is approximately the thickness of the gold plating.

solution in which the components diffuse at different rates. If the gold diffused into the nickel-iron more rapidly than nickel and iron diffused into the gold, a condition of "supersaturated" vacancies could arise in the gold layer. Bardeen and Herring,¹⁴ and Seitz¹⁵ have discussed a mechanism explaining unequal diffusion which assumes a substitutional exchange of atoms and vacancies with higher probability for one kind of atom than for the other. This process results in a portion of the current of atoms of the more rapidly diffusing component being balanced by a flow of vacancies in the opposite direction. Baluffi and Alexander attribute the formation of pores to the precipitation of these vacancies.

A study of the cell size of Au-Ni alloys¹⁶ reveals that the curve of a_0 versus composition departs from linearity, particularly in the range 15 to 20 per cent Ni with a_0 just under 4 Å. This deviation is attributed to vacancy concentrations as high as 3.5 per cent. The preparation of the alloys¹⁶ involved prolonged heat treatment, and in consequence the vacancies are assumed to be in equilibrium with the lattice. In contrast, the conditions in our experiment are nonequilibrium ones, since in the equilibrium condition the gold layer would disappear by diffusion into the iron-nickel base.

Many of the problems studied in our electron microscope laboratory are initiated in other departments. The study is usually a cooperative undertaking. The author wishes to acknowledge his indebtedness to his colleagues for their discussions and suggestions, and in particular to K. H. Storks, H. Christensen, E. E. Thomas, and F. W. Vogel, in connection with the material here presented.

References

1. GABOR, D. 1944. *The Electron Microscope* : 30-37. Hulton Press. London, England.

2. ZWORYKIN, V. K., G. A. MORTON, E. G. RAMBERG, J. HILLIER & W. A. VANCE. 1945. Electron Optics and the Electron Microscope. Ch. 19 (esp. : 701-708). Wiley, New York, N. Y.
3. GLASER, W. 1951. Symposium on electron physics. Natl. Bur. Standards.
4. HILLIER, J. & E. G. RAMBERG. 1947. J. Applied Phys. **18**: 48.
5. CALBICK, C. J. 1951. Bell System Tech J. **30**: 798; Proc. ASTM to be published.
6. VOGEL, F. W., W. G. PFANN, H. E. COREY & E. E. THOMAS. 1953. Phys. Rev. **90**: 489.
7. DAC. ANDRADE, N. A. & J. G. MARTINDALE. 1934-5. Trans. Roy. Soc. **235**: 69.
8. SEITZ, F. 1952. Advances in Phys. **1**: 43.
9. GRIFFITH, A. A. 1920. Trans. Roy. Soc. **221**: 163.
10. Fracturing of Metals. 1948. Am. Soc. Metals Monograph. See, in particular, summary article : 200.
11. OROWAN, E. 1948-9. Repts. on Progress in Physics. **12**. 1950. Fatigue and fracture of metals. In MIT Symposium.
12. HOFFMAN, R. W., F. J. ANDERS & C. E. CRITTENDEN. 1953. J. Applied Phys. **24**: 231.
13. BALUFFI, R. W. & B. H. ALEXANDER. 1952. J. Applied Phys. **23**: 1237.
14. BARDEEN, J. & C. HERRING. 1951. Trans. Am. Soc. Metals. **43A**: 87.
15. SEITZ, F. 1950. Acta Cryst. **3**: 355; 1953. Acta Metallurgica. **1**: 355.
16. ELLWOOD, E. C. & K. Q. BAGLEY. 1952. J. Inst. Metals. **80**: 617.

THE ADSORPTION OF DYES TO MICROCRYSTALS OF SILVER HALIDE*

By W. West, B. H. Carroll, and D. L. Whitcomb

Kodak Research Laboratories, Eastman Kodak Company, Rochester, N. Y.

Introduction

The study of the adsorption of dyes to the surface of sparingly soluble salt crystals has long been a source of significant information on the general problems of adsorption and the nature of surfaces. Several years before the publication of Langmuir's fundamental contribution¹ emphasizing the importance of monolayer saturation of the surface by the adsorbate, isotherms were published by Marc² for the adsorption of several dyes on the surface of salt crystals, which showed a saturation plateau at high equilibrium concentrations of the free dye. In 1922, in a study of dye adsorption on the surface of sparingly soluble salts whose area was known by the radioactive-indicator method, Paneth and Vorwerk³ made the first estimate of the area of a dye molecule from the amount of dye present at saturation, interpreted as indicating the completion of an adsorbed monolayer, and thus initiated the dye-adsorption method of determining surface areas.

The adsorption to silver halide of the dyes used, under the designation of "optical sensitizers," to extend photographic sensitivity to longer wave lengths than those effective in undyed photographic emulsions naturally attracted attention in the study of optical sensitization. Sheppard and Crouch⁴ began this study in 1928 and, in this and subsequent investigations by Sheppard and his co-workers,^{5, 6, 7} some of the fundamentals were established for the adsorption of cyanine dyes to silver halides. With the help of calculations made by Huggins of the area of the dye molecules, it was concluded^{6, 7} that the saturation plateaus observed in the adsorption isotherms of a number of cyanines adsorbed to silver bromide microcrystals corresponded to the formation of a close-packed monolayer of essentially planar cations, oriented with the molecular planes steeply inclined to the surface in a configuration presenting the edge of the molecule to the surface (edge-on adsorption). The work of Sheppard and his collaborators has been confirmed in its essentials and extended by Davey,⁸ and by the authors^{9, 10} of the present paper, which contains a review of the current position on the adsorption of sensitizing dyes to silver halide systems, along with some hitherto unpublished observations.

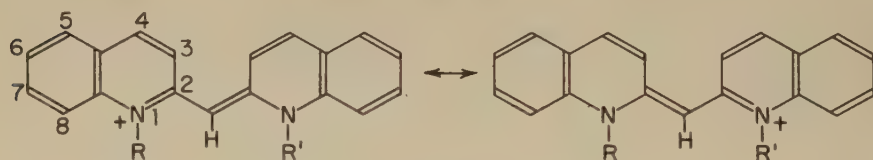
The Adsorption System

The adsorbent consists of the microcrystals of "grains" of silver halide of a photographic emulsion, usually crystallites of silver bromide containing a few per cent of silver iodide. The grains crystallize mostly in the octahedral habit of the cubic system; that is, (111) faces predominate. The average diameter of the crystals in a representative dispersion was about 0.8 μ . A normal photographic emulsion consists of a suspension of such microcrystals in an aqueous

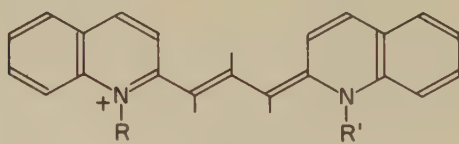
* Communication No. 1642 from the Kodak Research Laboratories.

medium containing approximately 5 to 7 per cent gelatin. The gelatin itself is known to be adsorbed to the surface of the grains, and an added dye therefore competes with gelatin for a place on the surface. The details of the adsorption of the dye are found to depend on the gelatin content of the suspension¹⁰ and, by varying the amount of gelatin, the activity of the silver halide surface can be varied with interesting modifications of the adsorption isotherms. The adsorption of dyes also depends on the concentration of silver ion or bromide ion in the medium. Unless otherwise stated, the data discussed will refer to suspensions approximately one-thousandth molar in bromide ion, at a temperature of 40° C.

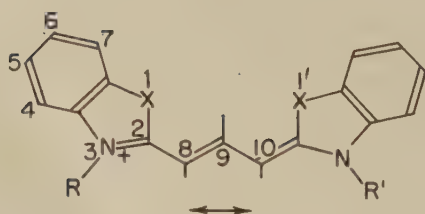
FIGURE 1 illustrates one type of dye used as adsorbate in this investigation, the cationic cyanine dyes. These consist essentially of a conjugated chain linking two basic nitrogen atoms on separate heterocyclic nuclei, such as quinoline, benzothiazole, benzoxazole, *etc.* The ions carry a single positive charge, which is not localized on any one atom in the conjugated chain, but which oscillates back and forth as a result of motions of the π -electrons of the chain.



2-2' Cyanine



2-2' Carbocyanine



$X = X' = O$, Oxacarbocyanine

$X = X' = S$, Thiocarbocyanine

$X = O, X' = S$, Oxathiocarbocyanine, *etc.*

FIGURE 1. Types of cyanine dyes.

It is well known how this type of structure is associated with intense absorption in the visible part of the spectrum. Another consequence of the nonlocalization of the charge in these molecules, first pointed out by London¹¹ and more recently discussed by Coulson and Davies,^{12, 13} is the abnormally high van der Waals forces emanating from these molecules. The intramolecular mobility of the π -electrons is equivalent to a high polarizability in the direction of the chain length, conducive to high values of the dispersion forces and, moreover, while the potential of the dispersion interaction between atoms or simple molecules varies inversely as the sixth power of the distance, the calculations^{11, 12, 13} show a slower diminution with distance of this interaction between long conjugated chains. The attractive forces between dye molecules are therefore strong and of relatively long range. Empirically, the tendency of dye molecules to dimerize and to form multimolecular aggregates, particularly in aqueous solution, is one of the most characteristic properties of these molecules. The strong van der Waals forces also, no doubt, contribute to the ready adsorption of dyes to crystalline surfaces, although electrostatic forces also play a part. But, as will be shown, the mutual attraction between dye molecules in the adsorbed layer also plays an important part and, on occasion, a decisive part in holding the adsorbed layer to the surface of the crystal.

The adsorption of certain merocyanines has also been studied. As shown in FIGURE 2, these dyes consist of heterocyclic nuclei joined by a conjugated chain,

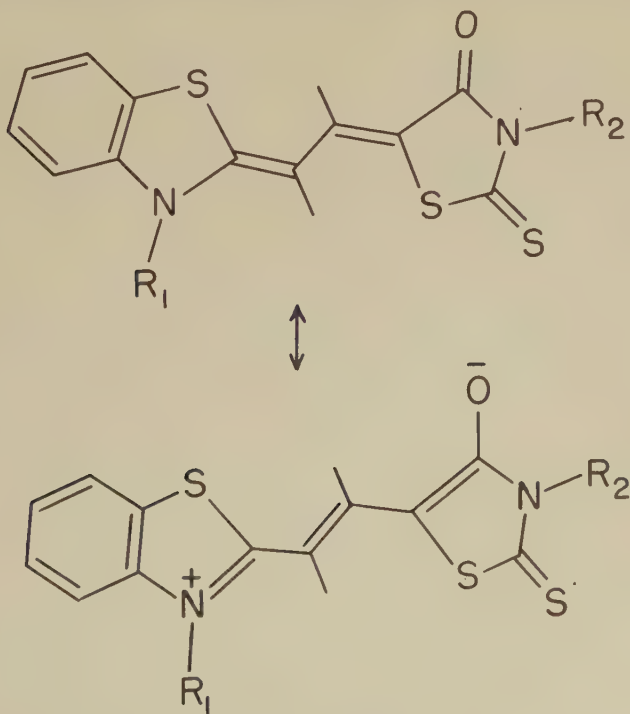


FIGURE 2. A merocyanine dye.

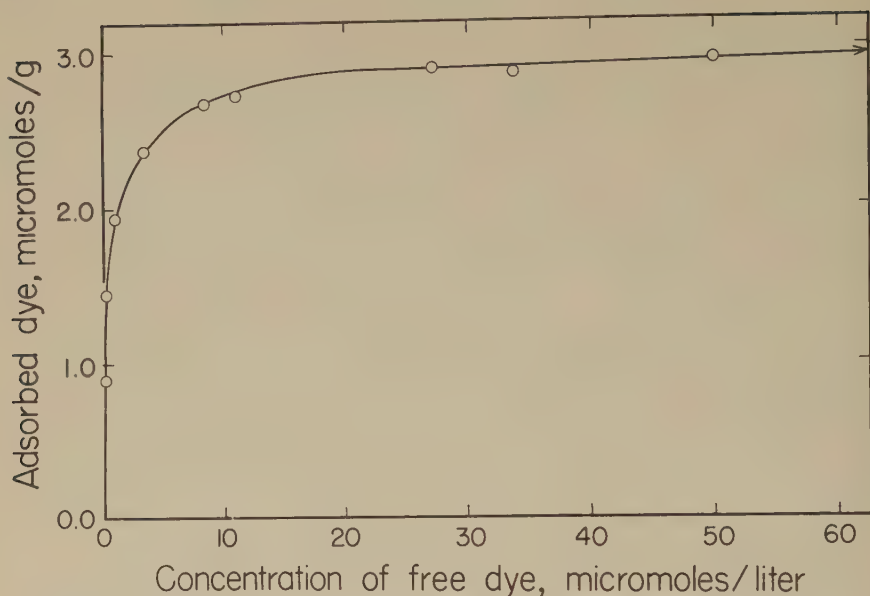


FIGURE 3. Langmuir-like adsorption isotherm of 1,1'-dimethyl-2,2'-cyanine iodide on silver bromide in 7 percent gelatin at 40° C.

one of the nuclei being basic, the other of acidic character. They bear no net charge but, as Brooker has shown, there is a nonlocalization of electronic charge of the same type as shown in the resonance of the amides.¹⁴ In terms of the theory of resonating structures, the merocyanines resonate between an uncharged and a dipolar structure, and the molecules, although neutral as a whole, have high dipole moments.

Langmuir-Type Adsorption Isotherms

FIGURE 3 illustrates what seems to be the simplest type of adsorption isotherm of dyes on emulsion grains. The obvious characteristic feature of this isotherm is the attainment of a saturation plateau. Sheppard's conclusion, already referred to,^{6, 7} that the saturation plateau represents complete monolayer coverage of the silver halide surface by dye cations in edge-on orientations, has been amply confirmed by our observations.⁹ To give a specific example, FIGURE 4 contains a scale drawing of the 1,1'-dimethyl-2,2'-cyanine cation. With the exception of the hydrogen atoms on the methyl groups attached to the N-atoms, it is very probable that all the atoms in this molecule are coplanar, the trigonal hybridization of the wave functions of carbon atoms in the conjugated chains tending to secure this arrangement. The smaller circles represent the van der Waals radius of the hydrogen atoms, and the larger circles similarly represent that of the methyl groups. The length of the molecule is 15.9 Å, its height is about 8.1 Å and, therefore, its area in flat orientation is about 130 Å². The minimum thickness of the molecule will be the van der Waals thickness of the benzene ring, 3.7 Å. The minimum area, which would

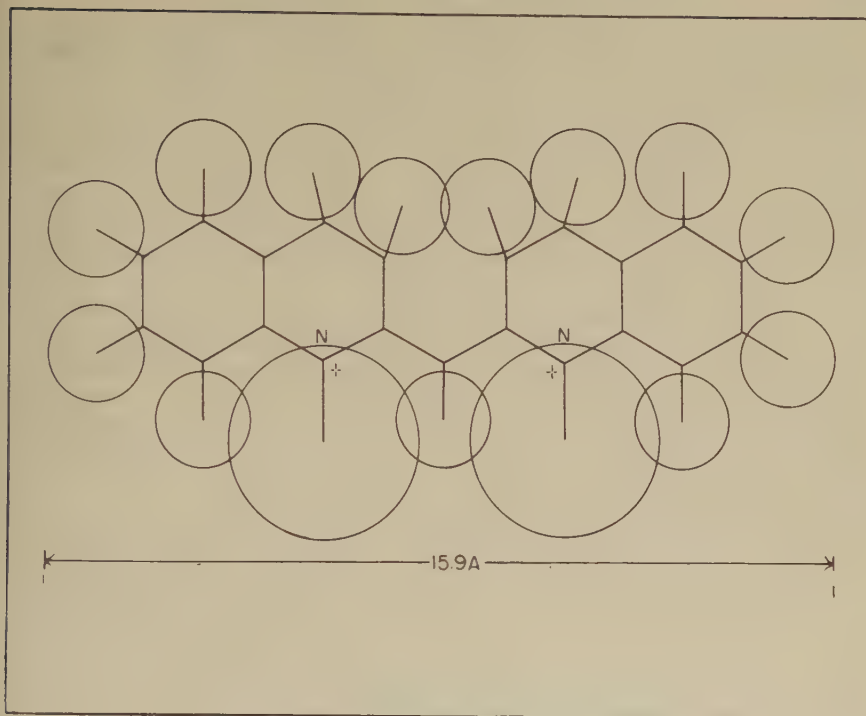


FIGURE 4. Scale drawing of 1,1'-dimethyl-2,2'-cyanine cation.

be occupied by a molecule in edge-on orientation would, therefore, be $15.9 \times 3.7 = 59 \text{ \AA}^2$, and the minimum area for the improbable end-on orientation would be $30 \text{ \AA}^2$. The actual area occupied per molecule adsorbed at saturation is found by dividing the surface area by the number of adsorbed molecules. Microscopic observation gave a specific projective area of 4.03×10^3 square centimeters per gram of silver halide. The grains were mostly platelike triangles with a thickness about one fifth the side, or regular hexagons with a thickness about one half the side. The total area of such grains is about 3.3 times the projective area, or about 13.3×10^3 square centimeters per gram of silver halide. Dividing this area by the number of adsorbed dye molecules at saturation gives the value $73.6 \text{ \AA}^2$ for the area occupied per molecule. It is clear that, if the saturation plateau represents a single layer of dye molecules, with an average distance of $73.6/15.9 = 4.63 \text{ \AA}$ between the molecular planes, the orientation must be edge-on. If the molecules are adsorbed flat on the surface, then saturation corresponds to the completion of a double layer of adsorbed molecules. Such a proposal has recently been made by Ewing and Liu¹⁵ for certain dyes adsorbed on TiO_2 . Such an arrangement, however, seems improbable for the cyanine dyes on silver bromide.

The spectral evidence, to be discussed in a later paragraph, shows that the free dye in equilibrium with the adsorbed dye at saturation is usually in the

monomeric form, so that the adsorption process does not consist in the attachment of preformed dimers from solution. Dimerization might, of course, be facilitated at the surface, but the absorption spectrum of the adsorbed dye layer is frequently almost identical, not with the dimer spectrum, but with that of a multimolecular aggregate as it occurs in solution, which has been called the J-aggregate.^{16, 17, 18} Various observations suggest that this aggregate consists of a threadlike structure of molecules with their planes parallel and highly inclined to the axis of the thread. If this structure for the aggregate in solution is accepted and, if the similarity of the absorption spectrum of the adsorbed dye to that of the aggregate in solution is taken to indicate a similar structure in the surface layer, the observed surface area per molecule is consistent only with a monolayer of molecules in the edge-on orientation.

When multilayer adsorption occurs, as it sometimes does,⁹ it can be recognized by the continuation of the isotherm beyond the amounts of adsorbed dye corresponding to monolayer edge-on adsorption. The monolayer stage may be distinguishable as a definite step, but for a few very well adsorbed dyes, multilayers begin to be formed before the monolayer is completed, and no definite step appears.

It will be noted that for 1,1'-dimethyl-2,2'-cyanine, the interplanar distance between the dye cations at saturation, on the hypothesis of edge-on orientation, is 4.63 Å, a value greater than the van der Waals thickness of the benzene ring, about 3.7 Å. The average intermolecular distances of a number of cyanine dyes at saturation are found to be in the range 4.6 to 5 Å. It is, of course, possible that the excess of the average interplanar distance observed over the minimum for a close-packed array of molecules in edge-on orientation may indicate incomplete coverage of the silver halide surface at saturation. It seems unlikely that the whole surface is covered with a completely regular array of edge-on adsorbed cations with a common orientation of the molecular axes with respect to the surface, but rather that islands of aggregate commence independently from various centers on the surface, with different orientations of the molecular axes with respect to the surface. At saturation, it is probable that the meeting of the separate islands will leave a certain amount of dead surface, and the average intermolecular distance calculated from the total surface would exceed the true interplanar distance on a microscopic scale.

It is, however, very significant that the interplanar distances calculated from the saturation adsorption plateaus are very similar to those deduced by Hoppe^{19, 20} from X-ray observations of thin films of 2,2'-cyanine dyes made by drying down an alcoholic solution on a support. These films, many molecules thick, exhibited strong J-absorption bands, and yielded fiber diffraction patterns from which was deduced a structure of dye cations arranged along a fiber axis with their planes parallel to each other; the interplanar distance along the fiber axis was 4.5 Å (for 1,1'-diethyl-2,2'-cyanine), and the normal distance between the planes was 3.7 Å. The structure deduced by Hoppe, therefore, was an arrangement of close-packed planar molecules with the molecular planes tilted some 35° from the normal to the fiber axis. Other dyes of the 2,2'-cyanine series showed in the dry films similar structures with interplanar distances varying from 4.5 to 5 Å.

The conclusions on the nature of the adsorbed layer arrived at from the adsorption isotherm are in essential agreement with Hoppe's conclusions concerning the structure of the J-aggregate in thin films. FIGURE 5 illustrates the general way in which the adsorbed layer of dye seems held to the surface of the grain, although we do not know from adsorption data the details of the arrangement, for example, as to a linear or staggered disposition of the molecular planes about the axis of the aggregate.

For dyes which have been studied both in adsorption and by X-ray analysis, the intermolecular distance calculated from the saturation limit is approximately 4 to 8 per cent higher than Hoppe's value. This difference may be the result of dead surface, as already discussed. There is, however, some uncertainty in the determination of surface area from the projective area, owing to the shape factor, and it is not certain if the differences between the X-ray and adsorption values of the interplanar distances are really significant. At higher temperatures, or at silver-ion concentrations in the suspension medium differing from those for optimum adsorption, the saturation plateau of several dyes was below that appropriate to complete coverage at the normal interplanar distance, and there seems no doubt that practical saturation by dye can be attained in this system without completely covering the surface-parts of the surface, either because of the occupation of adsorption sites by competing adsorbate or because of low adsorption potential of the surface sites, may be inoperative for dye adsorption.

Absorption Spectra of Dye Aggregates

The relations between the absorption spectra of dye aggregates in solution and on the grain are so important with respect to the structure of the adsorbed layer that they should be briefly enlarged upon. FIGURE 6 shows the absorption spectrum of the cyanine dye, diethyl-2,2'-cyanine chloride, at various concentrations in aqueous solution. At low concentrations of the order of

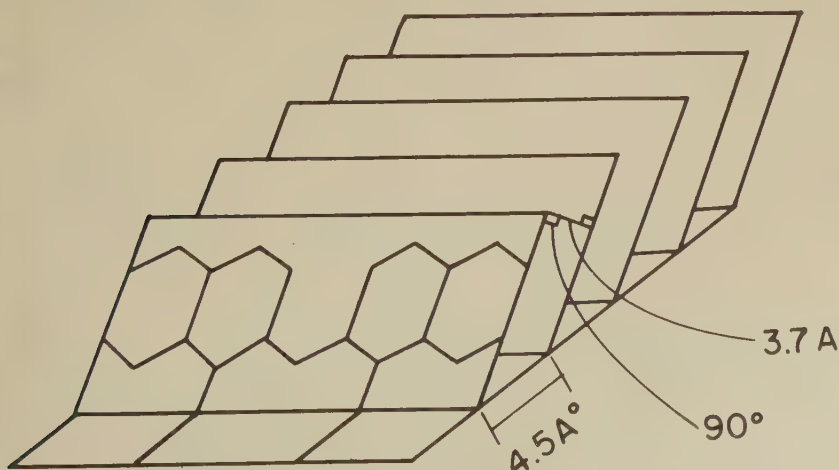


FIGURE 5. Arrangement of planar dye molecules adsorbed on silver bromide.

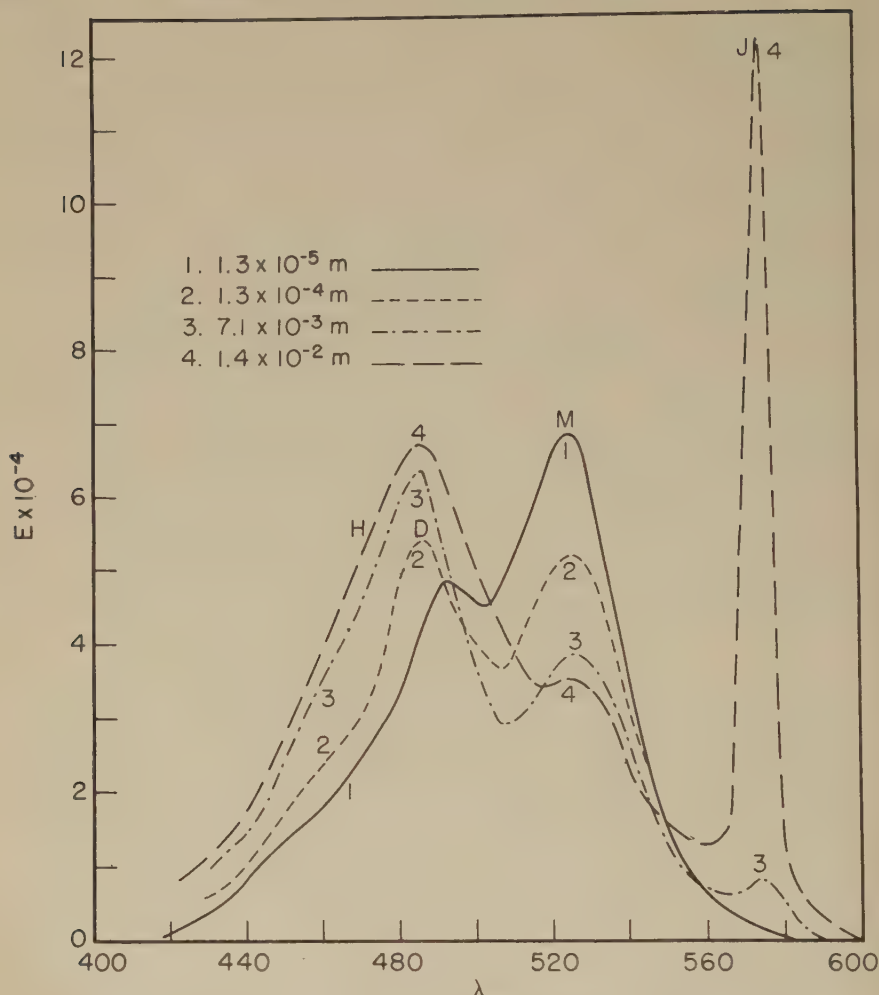


FIGURE 6. Absorption of aqueous solutions of 1,1'-diethyl-2,2'-cyanine chloride ($T = 25^\circ \text{C}.$).

10^{-5} molar, the spectrum is that of the monomer, very similar to the spectrum at all concentrations in alcoholic solution. The main features of the spectrum are a strong absorption maximum at about $522 \text{ m}\mu$ and a subsidiary maximum at $490 \text{ m}\mu$. With increasing concentration of the dye in aqueous solution, a band appears at a wave length somewhat shorter than the subsidiary band of the monomer, which seems quite independent of the latter band. This new band in aqueous solution seems attributable to a dimer.* At still higher con-

* For 1,1'-diethyl-2,2'-cyanine and for many other dyes, the steady passage of the dimer to higher aggregates with overlapping bands as the concentration of dye in aqueous solutions is increased makes it very difficult to show the presence of the dimer by deducing a dimeric equilibrium constant, using the spectral data to determine the concentrations of monomer and dimer. For a few dyes, however, the complication of further aggregation of the dimer is absent over a wide range of concentrations and, for these, as will be reported elsewhere, satisfactory dimeric equilibrium constants are found. It seems, therefore, safe to conclude that, in general, for cyanine and merocyanine dyes, the first change in the molecular species as the concentration of dye in aqueous solution is increased is the formation of a dimer.

centrations, the dimeric band becomes broader and is displaced toward shorter wave lengths, and it seems likely that the absorbing species is some aggregate higher than the dimer, termed an H-aggregate.²¹ Up to this point, the changes described are general for a great variety of dyes, cyanines, merocyanines, oxazines, thiazines, triphenylmethane dyes, and others. The next stage, however, the formation of the extraordinarily sharp band, designated J- in the figure, bathochromic to the monomeric band, is of rarer occurrence. This type of spectrum was first observed for this dye by Jelley,¹⁶ and at about the same time, by Scheibe.¹⁷ From a study of the variation of the intensity of absorption in this band with the direction of polarization of the incident light, these authors concluded that the absorbing entity was an aggregate consisting of a linear structure of dye molecules lined up with their planes parallel and highly inclined to the long axis of the chain—in fact, essentially the arrangement suggested by the adsorption isotherm and by Hoppe's X-ray observations. A later study of the J-band in solution was made by Mattoon,²² which emphasized the importance of water in determining the position of the J-maximum.

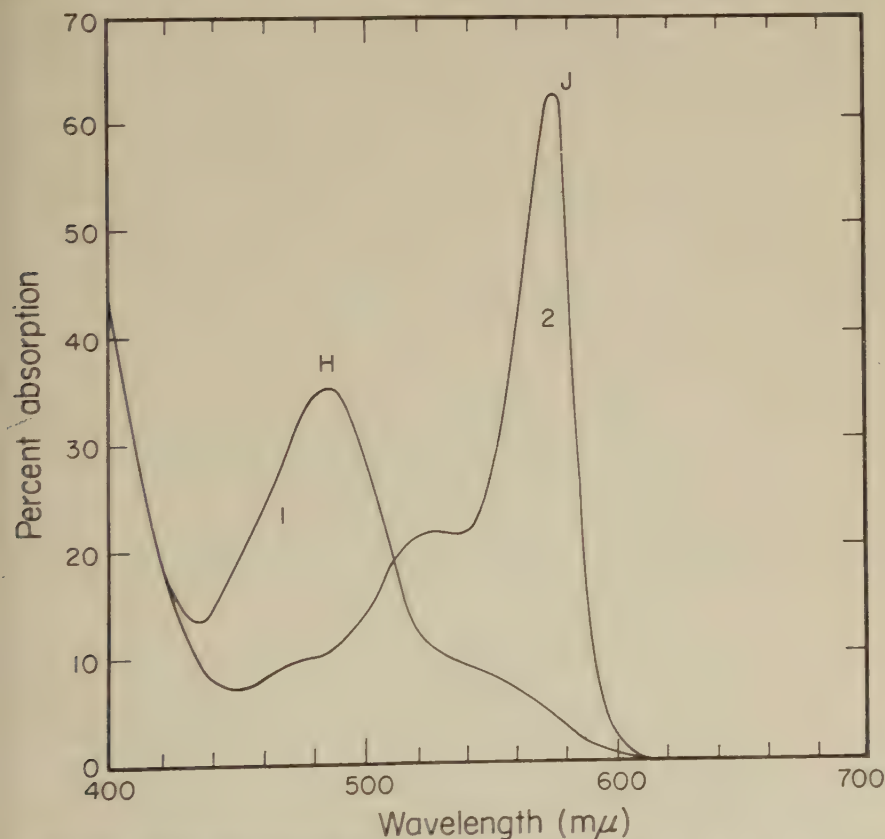


FIGURE 7. Typical H- and J-absorption bands of cyanine dyes adsorbed to silver halide. (1) 1,1'-Dimethyl-2,2'-cyanine iodide, showing strong H-band. (2) 1,1'-Diethyl-2,2'-cyanine iodide, showing strong J-band.

The spectrum of the adsorbed dye traverses a similar series of changes with increasing concentration of added dye. At the lowest concentrations, the spectrum consists of a broad band, with a maximum about 200 to 300 Å longer than that of the monomer in solution, which is very probably that of the isolated adsorbed molecule, the bathochromic shift indicating a tighter binding of the molecule to the surface in its excited state than in the ground state. Increase in concentration leads to short-wave displacements of the maximum, to positions similar to those of the H-aggregates in solution and, finally, for a limited number of dyes, strong J-maxima appear, at wave lengths near those in the corresponding solution.²³ FIGURE 7, curve 2, illustrates the absorption spectrum of 1,1'-diethyl-2,2'-cyanine iodide adsorbed to silver halide grains at approximately 40 per cent coverage. There can scarcely be any doubt that the structure responsible for this absorption is essentially the same as that associated with the J-band in solution. The concentration of added dye at which the J-band appears in the adsorbed layer is much lower than that at which it appears in solution. Aggregates are formed in the adsorbed layer more readily than in solution.

Not every dye adsorbed in edge-on orientation as deduced from the limiting plateau of the adsorption isotherm exhibits a J-band in the adsorbed condition. For example, 1,1'-dimethyl-2,2'-cyanine iodide adsorbed to silver bromide exhibits a large proportion of H-aggregate and, in some emulsions, is present mostly in that form. Curve 1 of FIGURE 7 illustrates the adsorption of this dye in a chlorobromide emulsion. No information from X-ray data and little from polarization measurements is available on the H-aggregate, and, although there are speculations in the literature on the difference between H- and J-aggregates,^{24, 25} this matter is at present little understood. Whatever these differences may be, the molecular units in both types of aggregate seem to be adsorbed to silver halide in an edge-on orientation, with approximately the same intermolecular distance.

Determination of Surface Areas by Dye Adsorption

An obvious application of the well-marked saturation plateau corresponding to monolayer coverage is the determination of surface areas of insoluble crystalline powders. Our own studies have been limited to silver halides, with and without gelatin. TABLE 1 contains data for the surfaces of three different silver halide suspensions, in gelatin, as estimated from microscopic analysis and from the limiting adsorption of 1,1'-diethyl-2,2'-cyanine iodide. The

TABLE 1
COMPARISON BETWEEN SURFACE AREAS ESTIMATED FROM MICROSCOPIC ANALYSIS AND FROM DYE ADSORPTION

Emulsion grains	Specific surface from microscopic analysis cm. ² /gram	Specific surface from dye adsorption cm. ² /gram
Bromiodide.....	13.3×10^3	13.3×10^3
Chlorobromide.....	18.5×10^3	19.5×10^3
Bromide.....	25.8×10^3	24.7×10^3

data for the bromoiodide are used to estimate the area occupied by the individual dye molecule, and this value of area of the dye is used to compute the areas of the other emulsions. Since the total area estimated from microscopic determination of the average projective area of a grain depends on a shape factor which is subject to error of about 5 per cent, the agreement between the two modes of estimation is satisfactory.

One requirement for the absolute estimation of total surface area by the adsorption method is that the saturation amount of dye be independent of temperature. This requirement is fulfilled for 1,1'-diethyl-2,2'-cyanine iodide and other dyes equally strongly adsorbed, for which the saturation plateaus were the same at 30°, 40°, and 50° C. The saturation amount of less well adsorbed dyes becomes progressively lower with increasing temperature, suggesting the presence of areas inactive in adsorption.

Cooperative Adsorption

Considerable light on the adsorption process of dyes on crystalline surfaces is cast by a study of the type of isotherm shown in FIGURE 9, curve 2. Characteristic of this curve is a sharp break, between a region of low adsorbability at low concentrations of added dye and a region of strong adsorption at higher

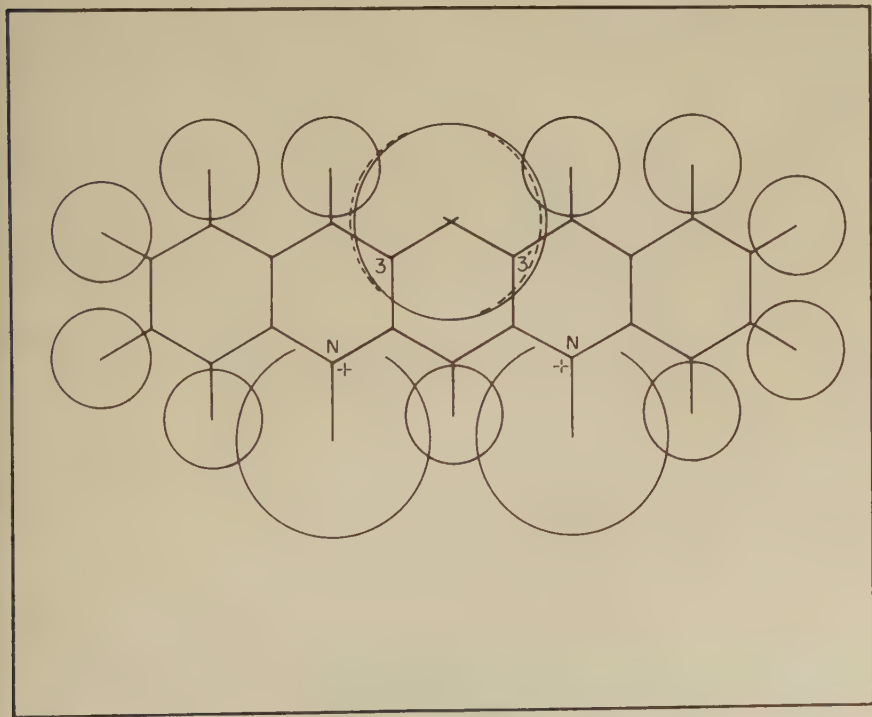


FIGURE 8. Scale drawing of 1,1',3,3'-tetramethyl-2,2'-cyanine cation, showing crowding of groups in the 3- and 3'-positions.

concentrations. The break is often abrupt, and occasionally there is practically no adsorption at all below a critical concentration, beyond which the isotherm rises rapidly to a limiting concentration corresponding again to edge-on adsorption.

This type of curve is frequently shown by less well-adsorbed cyanines. Relatively high solubility in water is one of the factors which may be associated with this type of isotherm, as well as environmental factors, such as the silver- or gelatin content of the dispersion medium, tending to decrease the activity of the adsorbing surface. An important structural characteristic associated with poor adsorbability is nonplanarity of the molecule. FIGURE 8 represents a 2,2'-cyanine containing the same nuclei as the well-adsorbed dyes of this class which have already been discussed, but with the substitution of methyl groups for the hydrogen atoms in the 3- and 3'-positions. The bulkiness of these groups forces the two nuclei from coplanarity, and the molecule must adapt a configuration in which the heterocyclic nuclei are twisted about the methine bridge. Such a configuration causes diminished overlap of the π -orbitals in the conjugated chain. There is less delocalization of charge, and the various properties depending on charge delocalization, such as intensity of absorption and the strength of the van der Waals forces leading to association and adsorption, are less marked. Curve 2 of FIGURE 9 shows the low adsorbability of the tetramethyl dye compared with that of its planar counterpart

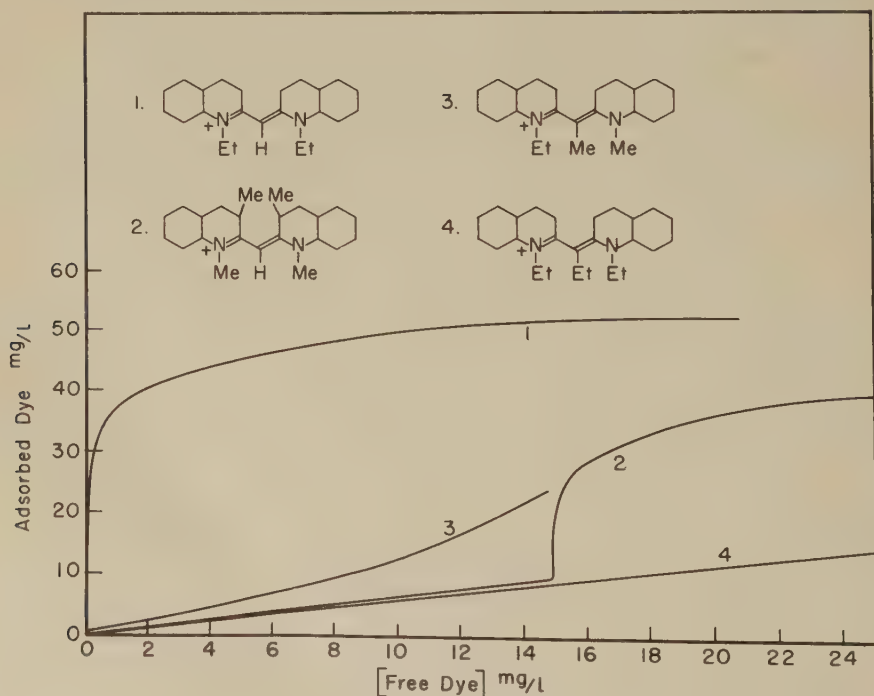


FIGURE 9. Effect of nonplanarity on adsorption.

(curve 1), while curves 3 and 4 show the low adsorption of 2,2'-cyanines forced from planarity by bulky substituents in the methine bridge.

One of the most interesting features of the absorption of dyes exhibiting the discontinuous type of isotherm is that the discontinuity is accompanied by a change in the absorption spectrum of the adsorbed dye. FIGURE 10 illustrates this for the nonplanar 1,1',3,3'-tetramethyl-2,2'-cyanine. In the low-con-

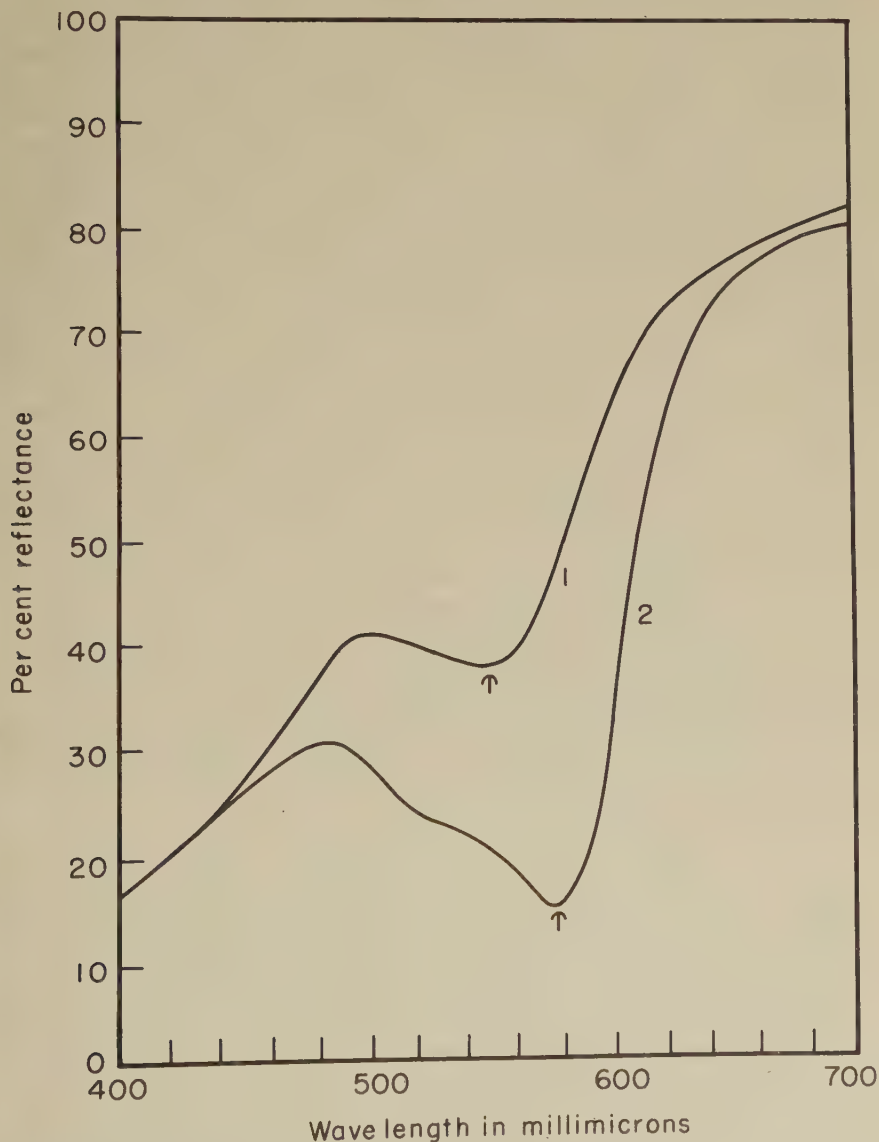


FIGURE 10. Reflectance spectrum of a nonplanar cyanine (1) in noncooperative adsorption (2) in cooperative adsorption.

centration foot of the isotherm, the spectrum is that of the isolated molecule as modified by its adsorption to the surface, while in the other region, the maximum undergoes a bathochromic shift to a wave length similar to that of the J-band in the corresponding planar dye. Although the new band is somewhat less sharp than the most typical J-band, there seems little doubt that it indicates a structure of the adsorbed layer beyond the adsorption discontinuity essentially similar to the oriented aggregates of planar dyes. For the non-planar dye, if the adsorption plateau is identified with the completion of a monolayer, the area per molecule is consistent only with edge-on adsorption, although the average intermolecular distance, 5.98 Å, is greater than for the corresponding planar molecules, as might be expected from the twisted configuration of the molecule.

The picture, then, which we evolve of the adsorption of dyes which exhibit this type of isotherm, is that at low concentrations the dye is adsorbed as isolated molecules, quite possibly in flat orientation with respect to the crystal surface. At a critical concentration, lateral van der Waals interactions between the adsorbed molecules become strong, and a cooperative layer in edge-on orientation begins to form. Once the layer begins to form, further addition of molecules from solution readily occurs, and the oriented layer rapidly covers the surface.

For dyes such as 1,1'-dimethyl-2,2'-cyanine, whose isotherm is reproduced in FIGURE 3, the onset of cooperation occurs at very low concentrations, and the isotherms on the scale shown appear continuous. Nevertheless, at very low concentrations, a small noncooperative foot can sometimes be found in the isotherm of well-adsorbed dyes which can be accentuated by choosing less favorable adsorption conditions, as by introducing competing adsorbate or an unfavorable silver-ion concentration.

Occasionally, adsorption is practically conditioned by cooperation. There is then practically no adsorption below a critical concentration at which the isotherm suddenly springs up steeply from the axis of abscissas. In the observed cases of this behavior, represented by the adsorption of a group of merocyanines containing the pseudindol and pyrazolone nuclei, a saturation plateau is attained at much less than complete coverage of the surface, only a small portion of the surface of bromiodide crystals being sufficiently active to hold these dyes.⁹

Heats of Adsorption

Isosteric heats of adsorption of a number of cyanine and merocyanine dyes to emulsion grains have been determined from the 40° and 60° C. isotherms by means of the Clapeyron-Clausius equation.⁹ Of particular interest are some recent results¹⁰ on the heats of adsorption of a merocyanine containing the thiazoline and rhodanine nuclei, the latter ring being provided with a solubilizing substituent. The heat of adsorption of this dye, as of others, decreases with increasing gelatin content of the medium, indicating the necessity of displacing adsorbed gelatin from adsorption sites, but the most interesting feature is the rise in the heat of adsorption accompanying the onset of cooperation when the adsorbing grains are suspended in an aqueous medium contain-

TABLE 2
HEAT OF ADSORPTION OF A MEROCYANINE AT DIFFERENT COVERAGES

Fraction Covered (%)	Heat (Cal/Mol)
15 (Noncooperative)	8,020
20 (Noncooperative)	7,820
25 (Cooperative)	9,300
30 (Cooperative)	9,670
35 (Cooperative)	9,630
40 (Cooperative)	9,750
50 (Cooperative)	9,480
60 (Cooperative)	9,430

ing 0.16 per cent gelatin. Under these conditions, the division of the isotherm into noncooperative and cooperative branches is well shown. TABLE 2 contains the results. In the noncooperative branch, the heat of adsorption falls with increasing fraction of the surface covered, a normal occurrence possibly caused by the utilization of the most active sites by the first quantities of dye added. At the discontinuity, there is a sharp increase in the heat of adsorption, which begins to diminish again after about 40 per cent coverage. The increase represents a contribution to the total heat of adsorption by the lateral van der Waals interaction responsible for cooperative adsorption. Besides the well-known effects of attractive interaction observed by Orr²⁶ in the adsorption of argon to potassium chloride, other examples of this phenomenon have been recently furnished, also in the gas phase, in the adsorption of propyl chloride vapor on zinc phosphate by Rausch,²⁷ and in the adsorption of ethyl chloride on graphite by Pierce and Smith.²⁸ The present data show that the heat of cooperation can also be observed in adsorption from solution.

The Adsorption Forces

The role of dispersion forces in holding a cooperative aggregate together on the adsorbing surface has already been emphasized, and it is probable that such forces contribute significantly to the binding between dye molecules and the surface. The influence of the silver-ion concentration in the dispersion medium, however, shows the existence of an important electrostatic contribution to the adsorption forces.^{7, 9} The adsorption of cyanine cations in gelatin suspensions of silver bromide is relatively little affected by silver-ion concentration in the range between 10^{-6} and 10^{-9} molar, when the surface of the grains are negatively charged, but rapidly decreases at silver-ion concentrations above 10^{-5} molar, which is known to be on the positive side of the neutral point of silver bromide particles in the absence of gelatin. It seems highly probable, therefore, that the observed diminution in the adsorption of cyanine cations at high silver-ion concentrations is caused by electrostatic repulsion between the grain surface and the dye cation, but this fall in adsorption of cyanines at silver-ion concentrations on the positive side of the neutral point is not very sharp. Well-adsorbed cyanines can be present on a "silver body" surface, and cyanine dyes cannot be used as adsorption indicators. The dispersion forces between the dye ion and the silver bromide surface appear to be sufficiently large to overcome appreciable electrostatic repulsion.

Merocyanines are dipolar but of zero net charge, unless they contain solubilizing ionic groups. With the object of studying the effect of dipole moment on adsorption, the isotherms of the group of dyes indicated in FIGURE 11 were measured. The dye of highest dipole moment, 17.7 D, has been described by Kushner and Smyth²⁹ as possessing the largest dipole moment of any molecule other than a salt or a zwitterion. Dye 1 was the most tightly adsorbed of the

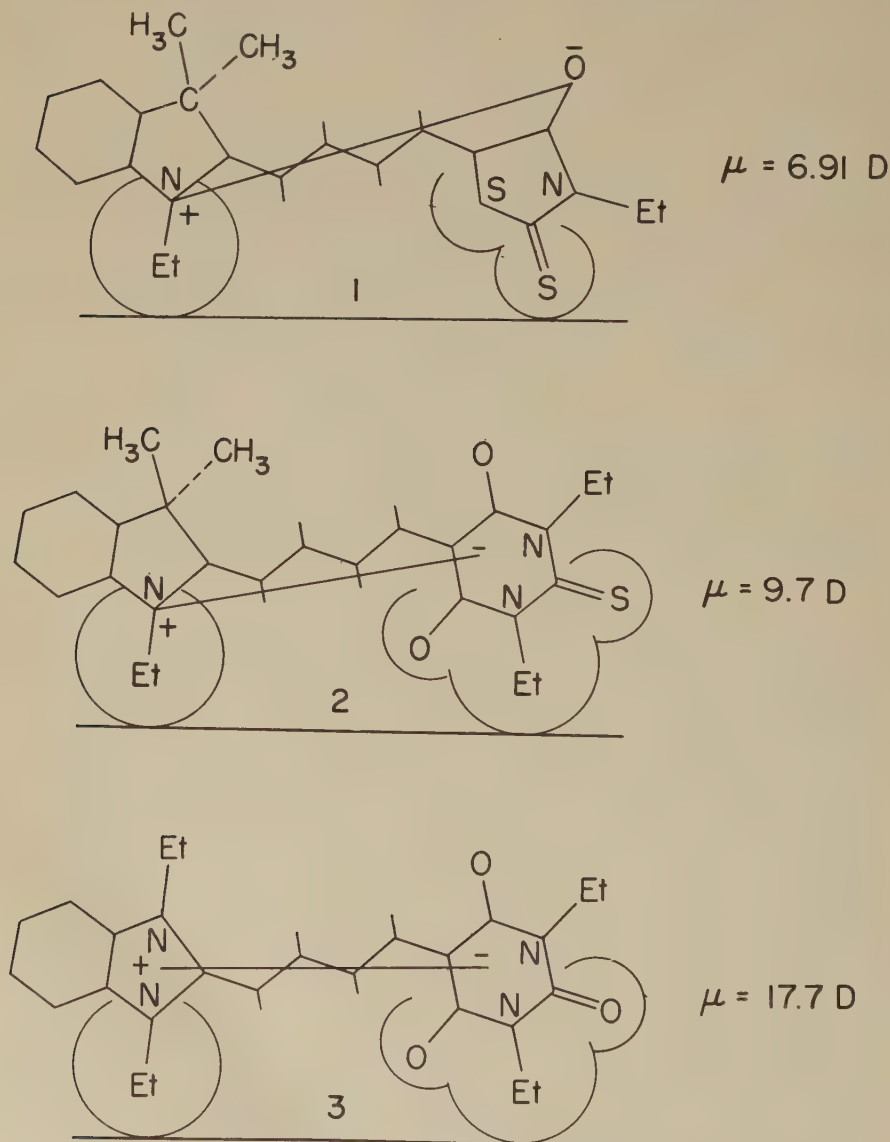


FIGURE 11. Merocyanines of high dipole moment.

group, in spite of its exhibiting the lowest dipole moment. Dye 3 was considerably more readily adsorbed than dye 2, which may possibly be the result of the high moment of the former molecule. It is clear, however, that among the merocyanines as a group, dipole interactions in themselves do not preponderate over other adsorption forces. The most tightly held dye in this group contains the rhodanine nucleus, with two sulfur atoms in strategic positions for edge-on adsorption to the surface, which the isotherm shows to be the orientation of the adsorbed layer. The adsorption of merocyanine dyes with sulfur-containing nuclei is found to be depressed by increasing the bromide-ion concentration of the dispersion medium, suggesting that these dyes are partly held by a sulfur-silver ion linkage of some sort. To sum up this somewhat fluid situation, it seems that interionic and possibly dipole forces contribute to the adsorption of cyanines and merocyanines to silver halide surfaces, dispersion forces play an important part in orienting the molecules of the adsorbed layer, favoring an edge-on disposition to the surface and contributing appreciably to the total binding energy, and that, for certain merocyanines, more specific forces of a quasi-coordination type play a decisive part.

Acknowledgment

The microscopic analyses of emulsion grains were carried out by R. P. Loveland, of these Laboratories, to whom our thanks are due.

References

1. LANGMUIR, I. 1918. *J. Am. Chem. Soc.* **40**: 1361.
2. MARC, R. 1911. *Z. physik. Chem.* **75**: 710.
3. PANETH, F. & W. VORWERK. 1922. *Z. physik. Chem.* **101**: 480.
4. SHEPPARD, S. E. & H. CROUCH. 1928. *J. Phys. Chem.* **32**: 751.
5. SHEPPARD, S. E., R. H. LAMBERT & R. L. KEENAN. 1932. *J. Phys. Chem.* **36**: 174.
6. SHEPPARD, S. E. 1938. *Atti Congr. intern. chim. 10th Congr. Rome.* 1938. **1**: 232.
7. SHEPPARD, S. E., R. H. LAMBERT & R. D. WALKER. 1938. *J. Chem. Phys.* **7**: 265.
8. DAVEY, E. P. 1940. *Trans. Faraday Soc.* **36**: 323.
9. WEST, W., B. H. CARROLL & D. L. WHITCOMB. 1952. *J. Phys. Chem.* **56**: 1054.
10. WEST, W., B. H. CARROLL & D. L. WHITCOMB. 1953. *J. Phot. Sci.* **1**: 145.
11. LONDON, F. 1942. *J. Phys. Chem.* **46**: 305.
12. COULSON, C. A. & P. L. DAVIES. 1952. *Trans. Faraday Soc.* **48**: 777.
13. DAVIES, P. L. 1952. *Trans. Faraday Soc.* **48**: 789.
14. BROOKER, L. G. S., G. H. KEYES, R. H. SPRAGUE, R. H. VANDYKE, E. VANLARE, G. VANZANDT, F. L. WHITE, H. W. J. CRESSMAN & S. DENT. 1951. *J. Am. Chem. Soc.* **73**: 5332.
15. EWING, W. W. & F. W. J. LIU. 1953. *J. Colloid Sci.* **8**: 204.
16. JELLEY, E. E. 1936. *Nature.* **138**: 1009.
17. SCHEIBE, G. 1937. *Angew. Chem.* **50**: 218.
18. SHEPPARD, S. E. 1942. *Revs. Modern Phys.* **14**: 303.
19. HOPPE, W. 1942. *Kolloid Z.* **101**: 300.
20. HOPPE, W. 1942. *Kolloid Z.* **109**: 21.
21. CARROLL, B. H. & W. WEST. 1951. *Fundamental Mechanisms of Photographic Sensitivity*: 169 *et seq.* J. W. Mitchell, Ed. Butterworths Scientific Publications Ltd. London, England.
22. MATTOON, R. W. 1944. *J. Chem. Phys.* **12**: 268.
23. LEERMAKERS, J. A., B. H. CARROLL & C. J. STAUD. 1937. *J. Chem. Phys.* **5**: 893.
24. FÖRSTER, T. H. 1946. *Naturwissenschaften.* **33**: 166.
25. SCHEIBE, G. 1948. *Z. Elektrochem.* **52**: 283.
26. ORR, W. J. C. 1939. *Proc. Roy. Soc. A.* **173**: 349.
27. RAUSCH, W. 1952. *Z. physik. Chem.* **201**: 32.
28. PIERCE, G. & R. N. SMITH. 1953. *J. Am. Chem. Soc.* **75**: 846.
29. KUSHNER, L. M. & C. P. SMYTH. 1949. *J. Am. Chem. Soc.* **71**: 1401.

ADSORPTION AND EXCHANGE IN METAL-METAL ION SYSTEMS

By Cecil V. King

Department of Chemistry, New York University, New York, N. Y.

Introduction

It is difficult to make quantitative measurements of the adsorption of metal salts from solution on metals, because the amounts adsorbed are so small. In classical methods, the solution must be analyzed and, to secure a satisfactory decrease in concentration, it is necessary to equilibrate a very large area of metal surface with a small volume of solution. This procedure normally involves the use of powdered or precipitated metal samples, with the attendant difficulty of obtaining accurate microscopic measurements of particle size and area. Then, there is no certainty that the true area is identical with the measured area. The roughness factor may be greater than unity because the powder does not consist of perfect crystals. Gas adsorption methods should probably be used to estimate the true area, since they seem more reliable than other methods; actually these methods have not been used.

Metals on which a film of oxide or other compound is stable when in contact with oxygen or water, of course, present another problem. A few investigators have treated such cases adequately, others have ignored the problem. The few easily available noble metals have received much attention. Platinum, however, is erratic in behavior, as is to be expected, since it absorbs and holds many substances so firmly.

If a finely divided metal is to be equilibrated with solution by stirring or rotating in a bottle, some solution is lost for the analysis, since it cannot all be drained away from the metal. In fact, it is not certain that forcible removal, as by pressure filtration or centrifugation, is desirable, since it may disturb the adsorption layer or the diffuse part of the electrical double layer, which may or may not be important in the adsorption process. If the metal powder is very fine, the necessary surface area can be secured without much bulk or much solution hold-up, but then the microscopic determination of area is less reliable. At a certain stage of coarseness, it is almost impossible to use a volume of solution which will wet the powder, yield enough for analysis, and show a satisfactory decrease in concentration.

One method of avoiding a large solution hold-up, which permits accurate determination of apparent surface area, is to employ a rolled-up metal sheet, which is fitted loosely into a glass tube. A small volume of solution can be washed over all parts of the surface repeatedly by shaking the tube.¹ In some respects, a better method is to percolate the solution through a column of finely divided metal. This method is similar to the chromatographic column technique, which has often been used with salt solutions and the usual adsorbents such as synthetic resins, soils, *etc.* The adsorption of copper sulfate on columns of finely divided copper and silver has been measured in this manner.² Instead of maintaining a head of liquid, the solution was dripped slowly onto the initially dry metal and allowed to percolate through without flooding the

column at any point. If equilibrium were attained at all points, the first effluent should be salt-free, and the concentration should rise very sharply to the initial value. In practice, because of mixing, channeling or slow attainment of equilibrium, the concentration rise is not so sharp. The total amount of adsorption, however, is accounted for by analysis of a comparatively small amount of solute-deficient liquid, in which the concentration change is relatively large.

At first glance, it would appear that the use of radioactive tracers would make adsorption measurements easy and quantitative. In practice, it is found that almost any solute deposits on almost any surface, often in amounts so large as to defy even qualitative explanation in terms of adsorption, exchange, roughness factor, displacement by impurities or local electrical cells, *etc.* A great deal of work will be necessary in careful preparation of clean surfaces of known area and a minimum of local potential differences and in establishing standard procedures, before quantitative explanations can be given.

The amount of exchange between a metal and its own ions is much greater than the amount of adsorption and most investigators have ignored the latter, washing the metal specimens thoroughly before measuring the radioactivity acquired from or lost to the solution. It is possible to measure both exchange and adsorption simultaneously, and experiments of this sort will be described below.

Measurement of Adsorption by Classical Methods

Silver salts on silver. The most extensive study of adsorption of its own ions by a metal has been made with silver and silver salts. The first measurements were made by von Euler and Hedelius,³ who employed reduced silver with an average particle diameter of 0.8 micron. They considered the particles as spheres rather than cubes, but this was probably of minor importance in calculating the surface area. One or two grams of silver was shaken with 15 ml. of silver nitrate solution for each measurement. The authors were not satisfied with their estimation of surface area, and new experiments were carried out by Rudberg and von Euler,¹ using silver sheets of 311 and 624 cm.² measured area, with only 5 ml. of solution. The amounts of adsorption found per cm.² apparent area differed considerably in the two cases, as will be shown below.

A few similar experiments with reduced silver were carried out by Tartar and Turinsky,⁴ who thought that any silver ion removed from solution was reduced by impurities remaining in the precipitated metal. In most of their runs, however, the ratio of solution volume to silver was too high to make significant measurements possible.

Our interest in this problem arose because of its importance in interpreting the rate of dissolution of silver in ferric salt solutions.^{5, 6} Silver ion, whether formed in the reaction or added separately, had a pronounced inhibiting effect, and rate equations could be based on adsorption isotherms of the Langmuir type. Adsorption measurements were carried out^{6, 7} with precipitated crystalline silver, which under the microscope was found to consist of uniform

rectangular parallelepipeds having a surface area of 500 to 675 cm.²/gram for different lots. Unfortunately, the effective surface area was not measured independently. Another lot of silver which had a microscopically determined area of 1780 cm.²/gram, was measured by low temperature nitrogen adsorption, which gave 2380 cm.²/gram.

Other measurements from which the amount of adsorption of silver nitrate on silver can be calculated have been made by Proskurnin and Frumkin.⁸ These authors were primarily interested in finding the concentration of silver ion at which there would be no potential difference at the silver-solution interface. This concentration was found at about 10^{-5} M silver nitrate. In one series of experiments the concentration change was measured, by an emf method, after a large silver sheet had been equilibrated with a small volume of solution. In another series, the total flow of electricity was measured between a large silver electrode which had been equilibrated with solution, and another which was freshly immersed in the same solution.

All of the adsorption measurements mentioned above fall into two groups: (a) on precipitated silver, except as mentioned below, the adsorption is always

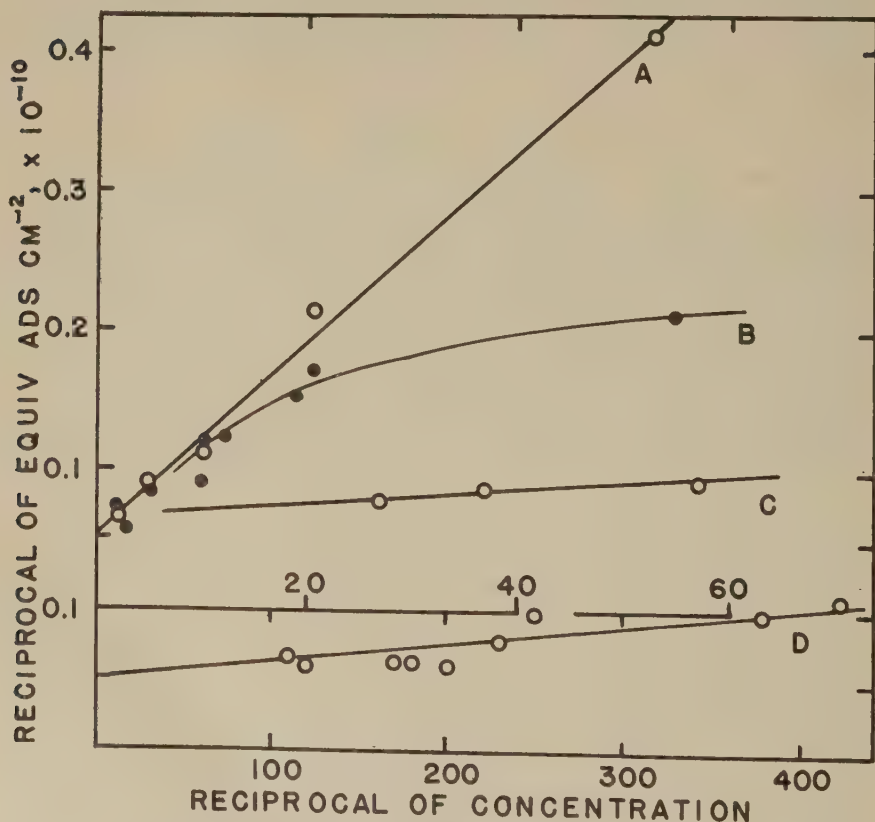


FIGURE 1. The adsorption of silver salts on silver. A, B, C: silver nitrate, perchlorate, and sulfate, respectively. King and Schochet. D: silver nitrate. Von Euler and Hedelius.

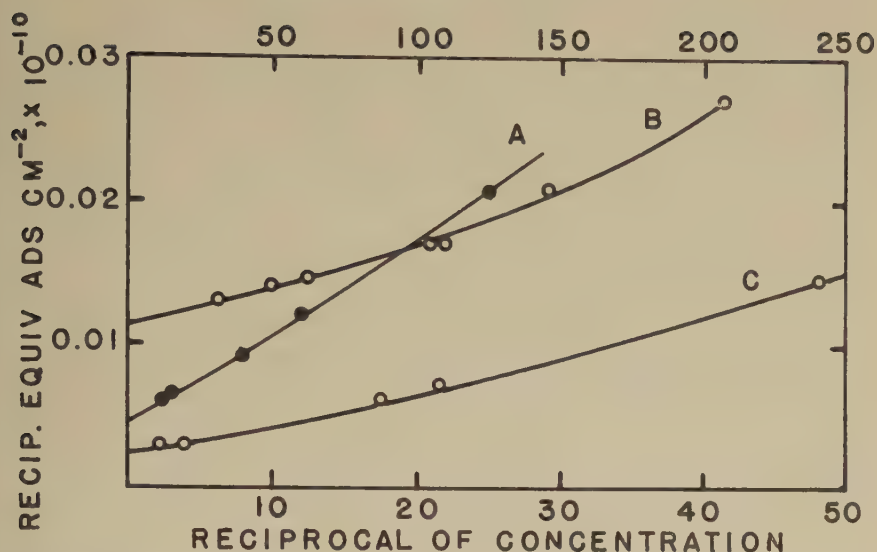


FIGURE 2. The adsorption of silver salts on silver. A: silver perchlorate on nitric acid-treated silver (upper scale). King and Lang. B: silver nitrate on sheet silver (upper scale). Rudberg and von Euler. C: silver perchlorate on nitric acid-treated silver (lower scale). King and Schochet.

less than monolayer in nature, and extrapolates to a value near monatomic layer coverage (based on apparent area), at higher concentrations; (b) the adsorption on nitric acid-treated silver, and on sheet silver in general, extrapolates to much higher values.

The measurements may conveniently be plotted as reciprocal of equivalents adsorbed per cm^2 apparent area versus reciprocal of solution concentration, as has been done in FIGURES 1 and 2. These plots would be linear if simple Langmuir isotherms were followed. In FIGURE 1, the values for the different silver salts are widely separated at low concentrations, and the two curves for silver nitrate do not agree at all. Yet it is remarkable that all the curves extrapolate, at $1/c = 0$, to about the same value, near 0.05×10^{10} . The extrapolated maximum adsorption is from 16 to 20×10^{-10} equivalents/ cm^2 (16 to 20 Gibbs⁹).

The reticular density in (100) planes of the silver crystal is 19.9×10^{-10} gram atoms/ cm^2 . It may be paralogistic, but it is tempting to conclude that the precipitated silver consists of nearly perfect crystals with clean surfaces and that, under the experimental conditions, the silver salt deposits with a monolayer as a maximum. The silver ions no doubt form a continuation of the normal lattice, and do not have to deposit uniformly, since their charges would not remain localized. The monolayer adsorption, if real, must be determined by a close association of anions with the surface. A small excess of positive charge must remain on the metal, to establish its potential with respect to the solution (at most a few tenths of a volt), and to account for the electrical double layer capacity, about 20 microfarads/ cm^2 .

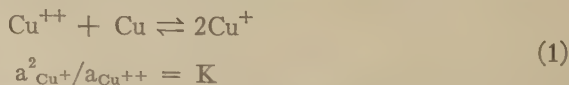
Treatment of the precipitated silver with 4.5 per cent nitric acid can increase

the amount of adsorption tremendously, as shown in curves A and C of FIGURE 2. This dilute nitric acid probably does not etch the surface or change the surface area appreciably, nor does it have a permanent effect; the same silver used to obtain curve A was later found to show adsorption comparable to the values in FIGURE 1. The sheet silver used by Rudberg and von Euler (curve B, FIGURE 2) was cleaned with steam and alcohol vapor. The extrapolated values in FIGURE 2 for maximum adsorption are $80, 200, \text{ and } 400 \times 10^{-10}$ equivalents/cm.² It might be possible that Rudberg and von Euler's sheet silver had a roughness factor of 4 but, in the other cases, it seems that multilayer adsorption must be assumed, unless the nitric acid leaves some material with reducing properties on the surface.

Proskurnin and Frumkin⁸ used silver sheet which was etched with nitric acid, washed, and later heated in hydrogen for each experiment. The amount of silver nitrate adsorbed varied from $20 \text{ to } 200 \times 10^{-10}$ equivalents/cm.², based on the apparent area, and depending on the concentration. Below 10^{-5} M silver nitrate, it would appear from the measurements, that appreciable amounts of silver actually dissolved (the solutions were deaerated and the vessel was evacuated). The amount of adsorption at equilibrium is possibly established more or less accidentally and dependent on the exact surface preparation. Starting with 1×10^{-4} M silver nitrate, Proskurnin and Frumkin found that 88 per cent of the salt was adsorbed (22×10^{-10} equivalents/cm.²), leaving a concentration of 1.12×10^{-5} M. On starting with 1×10^{-5} M silver nitrate, however, the concentration of silver ion apparently increased.

Silver salts on gold. The adsorption of silver salts on precipitated gold and gold leaf was studied by von Euler and Zimmerlund,¹⁰ and Rudberg and von Euler¹ carried out further experiments with leaf and sheet gold. Average values from the principal measurements are plotted as reciprocals in FIGURE 3. Adsorption values from extrapolation to $1/c = 0$ range from $70 \text{ to } 130 \times 10^{-10}$ equivalents/cm.² Since the lattice constants of gold are almost the same as those of silver, this range indicates either multilayer adsorption or roughness factors from 3.5 to 6.5. The precipitated gold particles had an average diameter of 1.2 microns, and could hardly have a roughness factor as high as 4.5 unless they were actually clusters of smaller particles, which seems unlikely. It is also improbable that the measurement of particle size would be so much in error. Multilayer adsorption is indicated.

Copper sulfate on copper and silver. In measuring the adsorption of copper ion on copper, a new complication is present. Copper (II) ion is in part reduced to copper (I) ion:



In practice, REACTION (1) can be ignored at concentrations greater than 10^{-3} M copper sulfate, but it becomes increasingly important in more dilute solutions. The constant K has for many years been considered to have a value near 10^{-6} , as determined in acidified solutions.^{11, 12} Newer measurements in unacidified solutions¹³ have indicated a value of about 10^{-5} . This equilibrium constant is

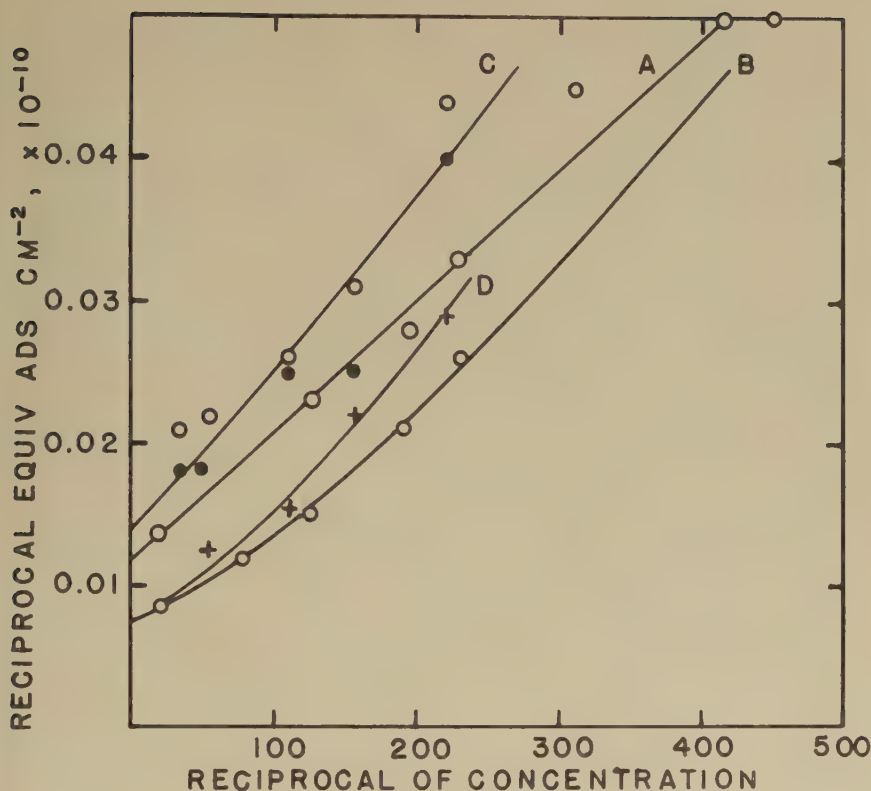


FIGURE 3. Adsorption of silver nitrate on gold. A, B: gold powder and foil, respectively. Von Euler and Zimmerlund. C: gold foil and gold sheet. Rudberg and von Euler. D: gold sheet, silver nitrate in 96 per cent alcohol.

important, for on it rests largely our knowledge of the Cu-Cu^+ and $\text{Cu}^+-\text{Cu}^{++}$ potentials, the free energy of formation of cuprous ions in solution, and the pH at which cuprous oxide will form on copper in contact with a cupric salt solution.¹⁴

When copper sulfate solution is shaken with copper in an atmosphere of nitrogen,¹³ or is passed through a column of finely divided copper,² the pH is lowered by 0.8 to 1.2 units, depending on the concentration. There is some question as to whether this decrease is entirely attributable to homogeneous hydrolysis of the cuprous sulfate formed, or partly to deposition of solid cuprous oxide on the copper surface. The magnitude of the change indicates the latter, but El Wakkad¹³ presents evidence to the contrary. In any case, if the original solution is acidified to a pH of 3 or 3.5, there is no appreciable change when in contact with copper.²

In the adsorption experiments carried out in this laboratory,² copper sulfate solutions were acidified, as above, and, after deaerating, were passed through a column of copper which had been heated in hydrogen and thereafter protected

from air. The amount of salt retained was in all cases equivalent to less than a monolayer, based on the apparent area. The roughness factor was estimated as from 3 to 5. Above 10^{-3} M there is fair agreement with either a Freundlich or a Langmuir isotherm.

At lower concentrations, the column effluent contains more total copper than the original solution because of REACTION (1). Actually, the final concentration had to be calculated in accordance with one or the other value of K . The adsorption of copper (II) and copper (I) ions cannot be calculated separately. The total adsorption falls off far more rapidly with dilution than simple adsorption theory can explain. This tendency is shown in FIGURE 4, where the vertical lines indicate the uncertainty in the value of K . The adsorption and the electrode potential probably fall to zero at about 10^{-6} M.

The adsorption of copper sulfate from methanol solutions, on copper, was found to be similar in magnitude. The same is true of the adsorption of copper sulfate from aqueous solutions on silver, if the relative true areas of the silver and copper were estimated approximately correctly. The adsorption on silver from methanol solutions is 12 to 15 times as great, and comparable to the amount of silver salts adsorbed on silver. Extrapolation, as before, to $1/c = 0$, gives about 16×10^{-10} gram ions of copper/cm.² on precipitated silver having an apparent area of 500 cm.²/gram.

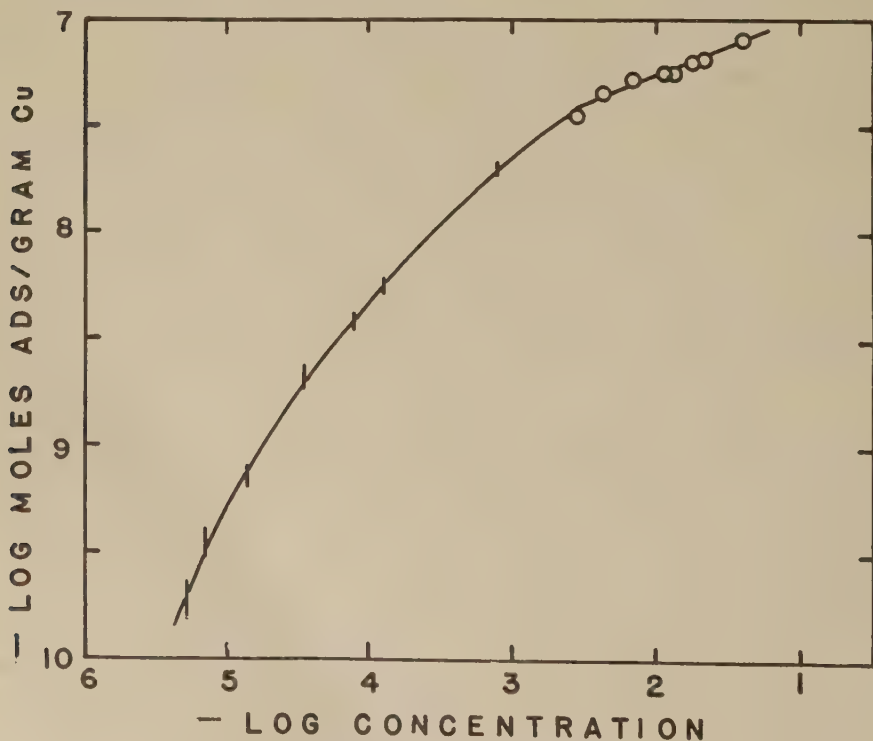


FIGURE 4. The adsorption of copper sulfate on copper.

Other metal-metal ion systems. Few other such systems have been studied, by classical methods, in more than qualitative fashion. In spite of the numerous studies of the mercury-salt solution interface, attempts to measure the amount of adsorption per unit area by direct methods have seldom been made. To obtain a large area, it has been necessary to use the falling-drop method. This gives little time for attainment of equilibrium, and requires an estimation of the volume which has been depleted in concentration, since the drops cannot be removed, as such, from the solution. Some experiments of this type were carried out by von Euler and Zimmerlund.¹⁰

Some measurements have been made of the adsorption of metal salts on finely divided platinum; for example, with lead and mercury salts, by Maxted.¹⁵ In such cases, the surface area has been totally unknown. The experiments are of interest because of the inhibiting effect on catalytic reactions. The adsorption of these salts on platinum must be specific and intense, since the surface becomes saturated at quite low concentrations. This is of course characteristic of substances which poison platinum as a catalyst.

Exchange at Metal Surfaces

It is not intended in this paper to review all of the studies of metal-metal ion exchange, but some of the work which is of interest in connection with adsorption will be discussed. It has already been mentioned that exchange may be far greater in extent than adsorption. For example, in some of the experiments with silver,⁷ radioactive Ag^{110} ion was added to the solutions as a tracer. After 24 hours, one or two per cent of the silver nitrate had been removed from solution by adsorption, while 60 to 80 per cent of the activity had been lost to the metal. This exchange was equivalent to replacing all the silver in 4 to 25 atomic layers on the crystal surfaces.

Theoretically, even less noble ions will deposit to a small extent on more noble metals, while the latter dissolve in equivalent amounts. This theorem is especially true if the two metals can form solid solutions. It is doubtful if such exchange has been detected and verified. In measuring the adsorption of copper sulfate from methanol on silver, it was noted that a small amount of copper could not be washed from the silver column with water. It could be removed by passing silver nitrate solution through the column. About 10^{-8} gram atoms of copper was thus recovered from 96 grams of silver with a surface area of 48,000 cm^2 . No silver ion could be detected in the effluent when copper sulfate solution was passed through the column, so it has to be assumed that this amount of copper ion was very firmly adsorbed and that no exchange took place. It has been assumed above, in general, that adsorbed salts could be removed fairly easily by washing, but this is certainly not always true. Other examples will be noted below.

Most of the published work on exchange has been qualitative in nature, experimentally and with respect to interpretation. A brief review of such experiments has been given by Simnad.¹⁶ A few attempts have been made, however, to obtain results worthy of quantitative interpretation.

It was first shown by von Hevesy¹⁷ that exchange takes place between a

metal and its own ions, using lead in lead nitrate solution, with Th B (Pb^{212}) as tracer. The exchange amounted in a very short time to the equivalent of 100 or more atomic layers. Similar unexpectedly large amounts of rapid exchange have been found in other systems. For example, with copper-copper salt solutions the exchange during an hour or less may be the equivalent of several thousand atomic layers.^{18, 19} The amount and rate depend on the concentration of the solution, and on the anions present.

These very large amounts of exchange customarily have been ascribed to the presence of local electrical cells on the metal surface, with resulting electrochemical dissolution at anodic areas and deposition at cathodic areas. Simnad¹⁶ and others have shown, by means of autoradiographs, that such cells are practically always present. Unfortunately, it is impossible to measure or calculate the emf of such cells or the amount of electrolysis for which they could be responsible, so that no quantitative theory can be formulated. If the metal specimen has been immersed in a solution of its own salt for a long time, it might be expected that such local cells would die away. However, it has been found (R. Glicksman, this laboratory) that when silver sheet was left in inactive silver nitrate for 5 months, then transferred to active solution for a few minutes, the autoradiographs still showed these irregularities.

Two excellent studies have been made of exchange in the silver-silver ion system. Gerischer and Vielstich²⁰ have measured both rates and magnitude of the process and have endeavored to interpret both. There is an initial rapid exchange which reaches a temporary "plateau" at a level of 20 to 50 "apparent" atomic layers, followed by a much slower exchange which continues indefinitely. The height of the initial plateau is lowered especially by pre-soaking in inactive silver salt solution. The time necessary to reach a definite fraction of the plateau height is inversely proportional to the concentration; with 0.1 M silver nitrate the plateau is reached in a few seconds.

Gerischer and Vielstich unfortunately did not make independent area measurements, but considered the roughness factor of their specimens to be 5 or 6, and that the initial rapid exchange took place in 3 to 5 atomic layers. They did not consider this to be a one-way process such as would be expected from local cell action, which they believed to be eliminated by pretreatment in inactive solution. A large part, though not all, of this initial exchange, is reversible. The authors postulated a two-way continuous exchange throughout several atomic layers owing to normal atomic fluctuations at the surface. The net result is equivalent to a diffusion coefficient downward through the surface layers, several powers of 10 greater than in the interior of the metal. The irreversible part of the initial exchange, and the slower exchange which continues indefinitely, was ascribed to interior diffusion.

Diffusion to the interior is, of course, a contributing factor, but most investigators disagree that it has much importance in the exchange at room temperature (*e.g.* E. Lange and others in the discussion of the Gerischer and Vielstich paper²⁰). The diffusion coefficient calculated by the authors from these measurements is several powers of 10 larger than is obtained by extrapolation to room temperature of the best work on silver lattice diffusion in the range 500–900° C.^{21, 22} It is at least several hundred times as high as the extrapolated

experimental values for polycrystalline silver,²¹ which are a combination of lattice and grain boundary diffusion; and about the same as the calculated value for pure grain boundary diffusion, or for surface diffusion.²³ The latter are thought to be similar in nature and in mechanism, and the high values are ascribed to a higher concentration of lattice vacancies than in the interior of the grains. It is not clear to what depth this rapid lateral diffusion can take place.

As Gerischer and Vielstich mention, experiments with single crystals would be very helpful.

The other outstanding study of silver-silver ion exchange has been made by Baerg and Winkler.²⁴ These authors made independent area measurements by a modified Bowden-Rideal method,²⁵ which involves electrodeposition of hydrogen on the metal with simultaneous measurement of the potential. Pretreatment in inactive solution did not change the surface area, which was about $15\times$ the planar area on emery abraded specimens, $7\times$ on etched samples. Pretreatment reduced the immediate exchange on the abraded metal to an apparent 13 to 15 atomic layers, while the corresponding exchange on etched metal was of the order of 3 apparent layers. The authors conclude that local cell action has a great deal to do with the process, and is especially important with the abraded metal.

Exchange equivalent to several hundred or several thousand atomic layers in a short time must surely be ascribed to electrochemical action, to surface recrystallization, or to the dissolution of impurities with irreversible deposition. Porous metals may have an apparently very high roughness factor (100 or more²⁶) but the common metals probably seldom have a roughness factor greater than 10. Bowden-Rideal measurements give higher values than other methods. Surfaces are compared with mercury or an amalgamated metal, and it is possible that the deposited hydrogen does not always merely cover the surface of other metals.

Of importance in the present report is the fact that Baerg and Winkler were led to the conclusion that silver ion is irreversibly or chemisorbed on the silver surface. In the Bowden-Rideal electrolysis, the potential did not rise normally after transferral from the exchange solution, but behaved as if silver ion were being discharged. This behavior was not eliminated by as much as one hour's washing with water. After the cathodic treatment, the specimen could very rapidly acquire additional radioactivity when reimmersed in the active solution, and the process could be repeated several times. Current required for discharging the silver ion on a 2.5 cm.² sample was about 550 microcoulombs, which amounts to an atomic layer based on planar area, or somewhat less based on true area.

If metal samples cannot be washed completely free of a certain portion of the adsorbed salts even by thorough washing, then previous methods and results will have to be reexamined.

Measurement of Adsorption and Area with Tracers

The removal of silver salts from silver by washing has been tested to some extent in this laboratory (B. Levy). Silver specimens were immersed in silver sulfate solutions containing radioactive S^{35} as tracer. In spite of the silver ion

exchange, only adsorbed salt will then be detected. With commercial rolled silver sheet, thorough washing with hot water removed most but not every trace of activity. With lightly abraded surfaces, considerably more activity remains.

To measure exchange and adsorption simultaneously, it is necessary, on removing the metal from the solution, to free it from practically all adhering liquid, count the activity, wash off all adsorbed salt, and count again. There are various ways of removing the liquid. Erbacher, in experiments to be described below, simply let his 8 x 8 mm. metal strips drain, then cut off and discarded the bottom half. This method has been tested on silver strips by B. Levy, using sodium sulfate, which is little adsorbed, with S^{35} as tracer. Commercial rolled sheet does not drain clean; nor does it give a "clean" test with a water spray method,²⁷ even after thorough pretreatment with grease-free organic liquids. After light abrasion, the metal responds to the spray test, but retains considerable activity after draining from the sodium sulfate solution.

Since the amount of adsorption will usually be greater than 10^{-10} gram ions/cm.,² let us assume that a counting precision equivalent to this figure will be satisfactory. An amount of 0.1 M solution equivalent to a film 10^{-6} cm. thick may then remain on the specimen. If the solution is more dilute, the adsorption will be smaller, but less solute will be present in such a film. Tests with sodium sulfate have shown that silver strips, rapidly and thoroughly blotted with filter paper or absorbent tissue after removal from solution, retain little more activity than corresponds to such a film.

TABLE 1 shows the results of experiments on the simultaneous measurement of exchange and adsorption of silver ion on silver using the above technique (R. Glicksman, this laboratory). The silver squares used were previously immersed in inactive 0.1 M silver nitrate solution for four months. The temporary plateau of exchange has already been reached in the first immersion time, and amounts to the equivalent of 2.5 to 3 atomic layers, based on the apparent area. The later slow rise is evident in the 22 hour runs. The amount of silver nitrate removed by washing corresponds to 0.3 to 0.5 atomic layers. The reason for the larger values at 22 hours is not known, but they may indicate that adsorption equilibrium has not been attained in the shorter times.

Other experiments with different pretreatment of the silver (as brief immersion in acidified thiourea solution) give quite different results, and no correlation with surface preparation has yet been made.

Erbacher²⁸ has carried out numerous experiments with lead, bismuth, and polonium ions (using Th B and Th C as tracers for the first two), and their adsorption and displacement reactions with nickel, silver, gold, and platinum. He has postulated that only adsorption occurs in certain cases, while in other cases there is monolayer electrochemical displacement plus adsorption on the deposited layer. He has evolved an elaborate method of estimating surface areas, based on these measurements.

For example, with bismuth ion on silver and lead ion on nickel in 12 per cent hydrochloric acid at 20°, with a stream of nitrogen, Erbacher finds only adsorption to take place, linear with concentration of the ion in question, and amounting to 0.1 per cent of the ion present in 1.5 ml. of solution. (This is

TABLE 1

ADSORPTION AND EXCHANGE ON SILVER IN 0.1 M SILVER NITRATE

(One inch squares of metal, a new sample for each immersion time. 620 cpm = 20×10^{-10} equivalent/cm.² Each double count refers to the two sides; counts are given to the nearest 25.)

Time of immersion	Counts per minute		
	Blotted	Washed	Difference
5 sec.	1700	1300	400
	1450	1100	350
10 sec.	1600	1250	350
	1500	1150	350
30 sec.	1650	1250	400
	1600	1200	400
60 sec.	1750	1400	350
	1500	1150	350
2 min.	1950	1700	250
	1800	1600	200
5 min.	1600	1400	200
	1850	1650	200
9 min.	1750	1550	200
	1750	1550	200
14 min.	2300	2000	300
	2050	1700	350
20 min.	1800	1600	200
	1950	1650	300
30 min.	1750	1600	150
	1750	1550	200
22 hrs.	3100	2450	650
	3000	2300	700
22 hrs.	3050	2200	850
	2800	2200	600

1500×10^{-10} moles, from 0.1 M solution, on the measured area of 1.377 cm.², correspondingly less from more dilute solutions.) With bismuth ion on nickel under the same conditions, Erbacher substracts from the total deposit 0.1 per cent of the bismuth ion remaining in solution, and finds that the remaining deposition reaches a value which is independent of concentration between 10^{-6} and 10^{-3} M bismuth ion. This amount, 50×10^{-10} moles on 1.377 cm.² he considers to be a monatomic layer of electrochemically deposited bismuth. Using an ionic diameter from other sources, he then calculates the effective area of the nickel to be 3.48 cm.² and the metal to have a roughness factor of about 2.5.

The area of gold sheet was estimated in the same manner, by deposition of

polonium from a solution containing 1 M hydrochloric acid and 0.9 M thiourea, and subtracting 0.1 per cent of the polonium ion remaining in solution as a correction for the amount adsorbed on the deposit. The roughness factor was found similar to that for nickel.

From more dilute (0.1 M) hydrochloric acid at 79°, for lead ion on silver and nickel, Erbacher finds a linear log-log relation between the amount adsorbed and the concentration. Polonium ion on gold at 20°, from 0.1 M hydrochloric acid, followed the same relation, which could be expressed by the equation

$$a = 8.3 \times 10^{-2} c^{0.65}$$

Here the units of a are 10^{-8} millimole and of c are 10^{-8} mole/liter. (This amounts to 320×10^{-10} moles adsorbed from 0.1 M solution, on the measured area of 1.377 cm.²) By using this equation as a correction factor for adsorption, the areas of nickel, silver, and gold specimens were estimated from experiments in these more dilute hydrochloric acid solutions. Roughness factors similar to the above were found, smaller on polishing with chalk, about the same on roughening with fine or coarse emery paper, *etc.*

Erbacher considers that hydrogen gas, adsorbed on the active portions of noble metals, may be electrochemically displaced by ions slightly more noble than hydrogen, such as lead and bismuth ions in suitable solutions. After correction of the total radioactive pick-up for the amount of adsorption as before, it is assumed that a monolayer remains, and that the active area may be calculated. In this way, the effective area of platinum, roughened with emery paper, was found to be 2.2, rather than 2.5, times the measured area. Erbacher takes this result to indicate that only 85 per cent of the surface is active. Polished platinum, heated to glowing, showed a much smaller active area.

Erbacher's method of surface area determination is obviously not suitable for general use, but it must not be dismissed lightly. The solutions were carefully chosen for each system, so that there was a small deposition potential. (For example, thiourea, by forming a stable complex ion, makes it possible for polonium ion to oxidize gold.) Ordinarily, electrochemical displacement does not stop at monolayer deposition, which requires an abrupt change in either the potential or the rate of deposition. The correction for adsorption on the deposit appears doubtful, but did not exceed 30 per cent, in the most useful cases 10 per cent, of the total deposit, because of the low concentrations actually employed for the electrochemical exchange. The actual potential of nickel in such solutions is in doubt, since the metal is in its passive state. More background for these experiments will be found in Erbacher's earlier papers, and brief discussions are given by Bonner and Kahn²⁹ and by Wang.³⁰

Monolayer adsorption. It has frequently been hopefully assumed that ionic adsorption on metals should end with monolayer coverage of the surface. This was done by von Euler and by Frumkin in the experiments described above. The authors were forced to conclude that either very close packing took place in the adsorbed layer, or that the available area of the metal was high, in some cases unreasonably high. Hackerman and Powers²⁶ conclude that chromate

ion adsorption on chromium does, indeed, stop at a monolayer. There is some evidence that irreversible chemisorption takes place, *e.g.*, the deposit is not removed by washing. The roughness factor of the porous electro-deposited chromium, as measured by krypton adsorption, was very high. The same authors³¹ find that somewhere near a monolayer of chromate ion is adsorbed on steel in a passivating solution. The exact value could be less than one or as high as 3.25, since the true area was in doubt. These specimens were rinsed, but not washed thoroughly. There was no adsorption unless the steel was passive.

Simnad¹⁶ has shown that the pickup of radioactive chromic and chromate ions from solution is extremely variable with different metals, depending on the pretreatment, and on the presence or absence of oxygen or oxide films; and in most cases is very uneven, as shown by autoradiographs. In the case of annealed copper as an example, the weight of chromium in the chromate picked up from deaerated solution (concentration not given) in 60 minutes was 1.76×10^{-6} gram/cm.². Simnad gives this as equivalent to 56 atomic layers; using the ionic area given by Hackerman and Powers,²⁶ it amounts to 100 layers. The specimens were washed, and it would seem that this activity was very firmly adsorbed, perhaps chemisorbed, or was in an insoluble product of reaction. The pickup of positive chromic ion was less than that of negative chromate ion.

It must be remembered that impurities or alloyed base metals may displace more noble ions. The appearance of iron in the effluent from a copper column with copper sulfate was noted,² although analysis of the metal showed iron to be present in very small amount. Haïssinsky, Cottin, and Varjabedian¹⁸ mention as specific examples, zinc in copper, silver or copper in gold. Many results with radioactive tracers, reported in the literature as adsorption, are probably due to some such type of electrochemical exchange.

Conclusions

We are forced to conclude that the study of adsorption and exchange at metal-solution interfaces is in its infancy, so far as precise experimentation and quantitative interpretation are concerned. The reasons are obvious. It is difficult to prepare and maintain clean metal surfaces, even to know when they are clean, or covered with a uniform film. The true area is usually unknown, or uncertain even when attempts are made to measure it. Little is known about the magnitude and potential of local electrical cells, or whether really equipotential surfaces can be prepared. Practically nothing is known about thermal or electrical atomic fluctuations at a metal surface in contact with a solution containing its own or other ions, or how such adjustments affect the reactivity in depth.

It is not really known whether the adsorption of salts on a metal resembles the adsorption of gases on solids, or is more like the deposition of additional layers on a salt crystal, from saturated solution. It is not known whether the adsorption is caused primarily by the requirements of the electrical potential and of building the electrical double layer, or is more specific in nature. It is

generally assumed that cations are primarily adsorbed on noble metals, dragging anions with them; but it is not known whether the reverse is true with base metals, or at what point in the series of metals the process may change.

There is no agreement in the literature as to what constitutes an adsorbed monolayer. Many authors base the calculation on an arbitrarily chosen area for the adsorbed ions; a few use the lattice parameter of the metal; many do not tell how the amount is calculated. This lack of agreement reflects, of course, the uncertainty as to how salts are adsorbed.

Partial answers to some of these problems have been given in specific cases, and it is evident that no easy or general solutions will quickly be found. Detailed study of each system is needed for the present, with great attention to surface preparation, true area determination, and local potential variations on the surface.

This report and some of the experimental work described in it were made possible by a research contract between New York University and the United States Atomic Energy Commission, to both of whom the author wishes to express his appreciation.

References

1. RUDBERG, E. G. & H. VON EULER. 1923. *Z. Physik.* **13**: 275.
2. SCHARFSTEIN, L. R. & C. V. KING. 1954. *J. Phys. Chem.* **58**: 180.
3. VON EULER, H. & A. HEDELIUS. 1920. *Arkiv Kemi. Mineral. Geol.* **7**(31).
4. TARTAR, H. V. & O. TURINSKY. 1932. *J. Am. Chem. Soc.* **54**: 580.
5. SALZBERG, H. & C. V. KING. 1950. *J. Electrochem. Soc.* **97**: 290.
6. KING, C. V. & F. LANG. 1952. *J. Electrochem. Soc.* **99**: 295.
7. KING, C. V. & R. K. SCHOCHET. 1953. *J. Phys. Chem.* **57**: 895.
8. PROSKURNIN, M. & A. FRUMKIN. 1931. *Z. physik. Chem.* **A155**: 29.
9. DEAN, R. B. 1951. *J. Phys. & Colloid Chem.* **55**: 611.
10. VON EULER, H. & G. ZIMMERLUND. 1921. *Arkiv Kemi. Mineral. Geol.* **8**(14).
11. FENWICK, F. 1926. *J. Am. Chem. Soc.* **48**: 860.
12. HEINERTH, E. 1931. *Z. Elektrochem.* **37**: 61.
13. EL WAKKAD, S. E. S. 1950. *J. Chem. Soc.* :3563.
14. POURBAIX, M. J. M. 1949. *The Thermodynamics of Dilute Aqueous Solutions.* Arnold, London, England. Ch. 6.
15. MAXTED, E. B. 1925. *J. Chem. Soc.* **127**: 73.
16. SIMNAD, M. T. 1952. *Monograph 13.* Institute of Metals. London, England. :23-58.
17. HEVESY, G. VON. 1915. *Physik. Z.* **16**: 52.
18. HAÏSSINSKY, M., M. COTTIN & B. VARJABEDIAN. 1948. *J. Chim. Phys.* **45**: 212.
19. QUINTIN, M., P. SUE & M. BIZOUARD. 1948. *Compt. rend.* **226**: 1723.
20. GERISCHER, H. & W. VIELSTICH. 1952. *Z. Electrochem.* **56**: 380.
21. HOFFMAN, R. E. & D. TURNBULL. 1951. *J. Applied Phys.* **22**: 634.
22. SLIFKIN, L., D. LAZARUS & T. TOMIZUKA. 1952. *J. Applied Phys.* **23**: 1032.
23. WINEGARD, W. C. & B. CHALMERS. 1952. *Can. J. Phys.* **30**: 422.
24. BAERG, A. & C. A. WINKLER. 1953. *Can. J. Chem.* **31**: 319, 521.
25. BOWDEN, F. P. & E. K. RIDEAL. 1928. *Proc. Roy. Soc.* **A120**: 59, 80.
26. HACKERMAN, N. & R. A. POWERS. 1953. *J. Phys. Chem.* **57**: 139.
27. LINFORD, H. B. & E. B. SAUBESTRE. 1951. *Plating.* **38**: 713, 847.
28. ERBACHER, O. 1933. *Z. Physik. Chem.* **A163**: 196, 215, 231; 1938. **A182**: 243, 256.
29. BONNER, N. A. & M. KAHN. 1951. *Radioactivity Applied to Chemistry.* 170. A. C. Wahl & N. A. Bonner, Eds. Wiley. New York, N. Y.
30. WANG, J. H. 1951. *Radioactivity Applied to Chemistry.* 316. A. C. Wahl & N. A. Bonner, Eds. Wiley. New York, N. Y.
31. POWERS, R. A. & N. HACKERMAN. 1953. *J. Electrochem. Soc.* **100**: 314.

THE ADSORPTION OF GASES AND VAPORS ON GERMANIUM

By J. T. Law and E. E. Francois

Bell Telephone Laboratories, Murray Hill, N. J.

1.0. Introduction

The importance of germanium in transistor physics has fostered the desire to know more about its nature. Electrical properties of germanium are markedly affected by surface contamination¹, and, in this connection, it was decided to ascertain, if possible, to what extent gases and vapors adsorb, the possibility being that information of this type might be correlated with the change in surface lifetime or contact potential under comparable conditions. It was hoped, also, that a better understanding might be obtained of the process involved.

The purpose of this paper is to present a survey of the adsorption properties of gases on germanium which have been investigated using a flash filament technique^{2, 3, 4} and a mass spectrometer as a pressure measuring device. Use of the mass spectrometer has the advantage that the composition of the gas desorbed can be determined and any reaction products identified. The adsorption of neon, argon, krypton, hydrogen, nitrogen, carbon monoxide, carbon dioxide, oxygen, water, and methyl alcohol have been studied in this manner.

2.0. Experimental Apparatus

Two filaments having the dimensions 2 mm. x 2 mm. x 100 mm. were cut from a single bar of zone-refined,⁵ high-resistivity, N-type germanium and molybdenum leads welded to the ends. The filaments were subjected to different treatments as follows:

2.1. On filament 1, the ends were masked with wax to protect the molybdenum leads and the rod treated with CP-4 etchant,* followed by thorough washing and removal of the wax with carbon tetrachloride.

2.2. Filament 2 was merely sandblasted and washed with distilled water.

The two filaments were mounted in glass jackets and connected to the inlet line of the consolidated mass spectrometer. FIGURE 1 shows a simplified schematic of the experimental arrangement. The heating current for the filament was supplied by means of a manually operated variac. Because of the large change in resistance of the filament in going from room temperature (10,000 ohms) to 200° C. (56 ohms), very little control of temperature over this range was possible. Above 200° C. the resistance decreases more slowly until at 800° C. it has a value of 0.3 ohms. During a flash run the mass spectrometer was set to focus on the mass of an ion fragment of the gas being investigated, a record of the gas released being made on the recording oscillograph. Temperatures within the range of visible radiation were determined with an optical pyrometer, and lower values were obtained by extrapolation, using the equation⁶ for the resistivity of intrinsic germanium. From these values a master curve of current versus temperature was constructed for each filament.

* This etch was developed by R. D. Heidenreich and has the following composition: 15 cc. of acetic acid, 25 cc. of concentrated nitric acid, 15 cc. of 48 per cent HF, and 0.3 cc. of bromine.

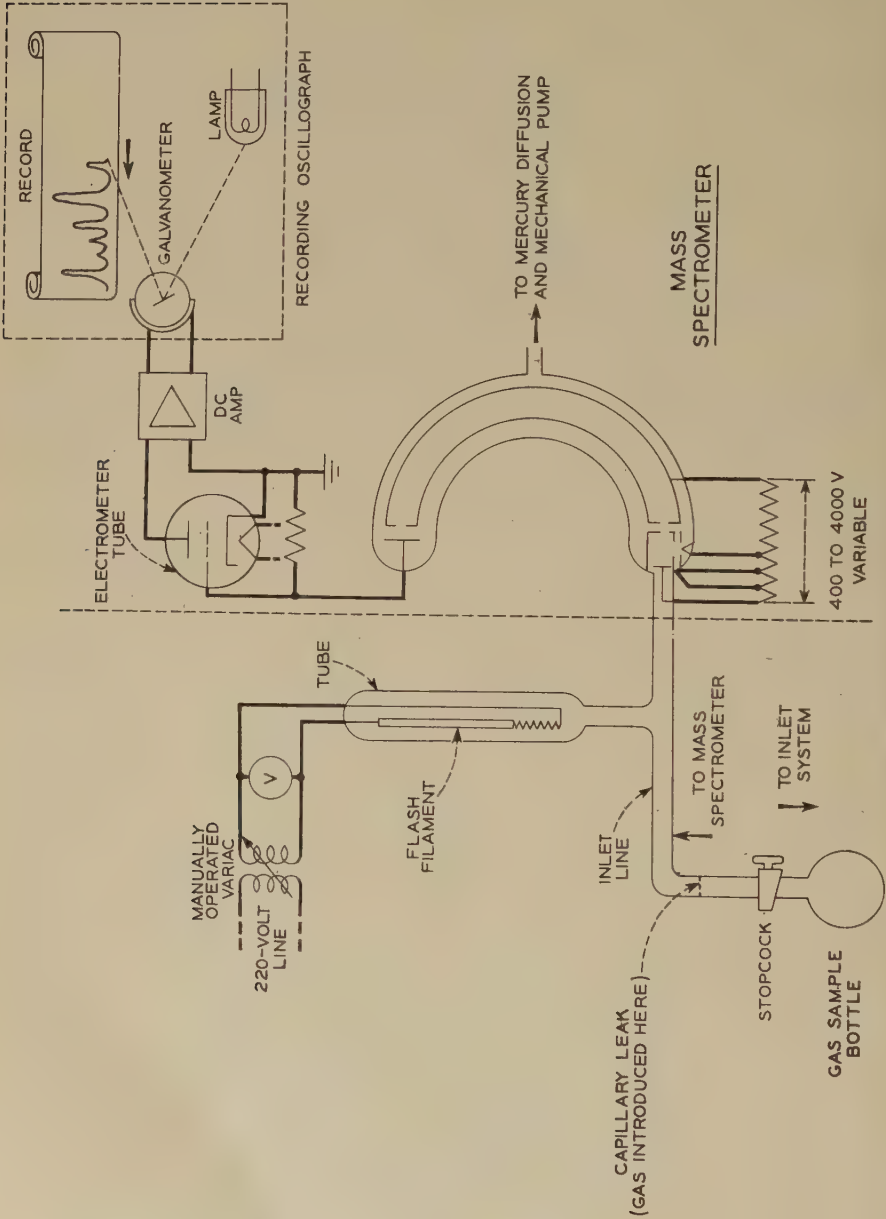


FIGURE 1. Experimental apparatus.

3.0. *Experimental Procedures*

The germanium filaments were outgassed before being subjected to any of the gases investigated. During this period, large quantities of water, hydrogen, carbon monoxide, and carbon dioxide were evolved. The carbon monoxide, carbon dioxide, and water vapor probably came from both the glass walls and from the germanium itself. In later experiments where the glass was first thoroughly degassed at 500° C., further appreciable amounts of gas were evolved when the germanium was heated to 850° C. The hydrogen was produced, at least in part, by decomposition of water at the germanium surface; this process will be discussed further in Section 5.5. Heating for extended periods at 850° C. finally eliminated most of these gases, and during the following experiments the temperature was never raised above 800° C. The methods used to investigate the adsorption and reaction of gases were as follows:

3.1. The filament was flashed three times to 800° C. and then the gas was introduced to the filament volume while the germanium was cold. After varying time intervals the gas was pumped out and the filament flashed, focusing on a peak of the gas. A typical resulting curve is shown in FIGURE 2, where the pressure of the gas being investigated is plotted against time. It may be seen that, when the germanium is heated to 800° C, the pressure at first rises sharply and then falls as the system pumps out. For the entire time shown beyond the initial pressure rise, the germanium is at 800° C. The peak heights obtained by flashing after pumping out the system will represent the amount of gas left on the surface. This amount will be only the same as that absorbed at the equilibrium pressure if the adsorption isotherm is not reversible. Whenever desorption occurs on lowering the pressure, values of the amount of gas adsorbed will be obtained which are lower than the true values at the equilibrium pressure.

3.2. The filament was heated to 800° C., and the gas was introduced while the germanium was hot. The filament was then cooled to room temperature in the presence of the gas, and the gas was pumped out. A flash run was made as before. For all the gases studied, no difference in peak height was found between the two techniques.

3.3. In cases where reaction occurred, for example with water, the following method was used: the mass spectrometer was focused on an ion fragment of one of the reaction products with the gas flowing through the system and the filament cold. The germanium was now heated in steps of 100° C. or 50° C. and a continuous record of the partial pressure of the decomposition product obtained. The reverse method of stepwise cooling from the maximum temperature was also used.

If the peak investigated was common to more than one gas it was necessary to focus on another peak that was peculiar only to the gas investigated in order to separate the effects. The purity of the gases used was checked by carrying out a sweep over the whole mass range in each case.

4.0. *Calculation of the Results*

The peak heights observed on the mass spectrometer represent pressure readings in the ion chamber and the relationship between peak height and

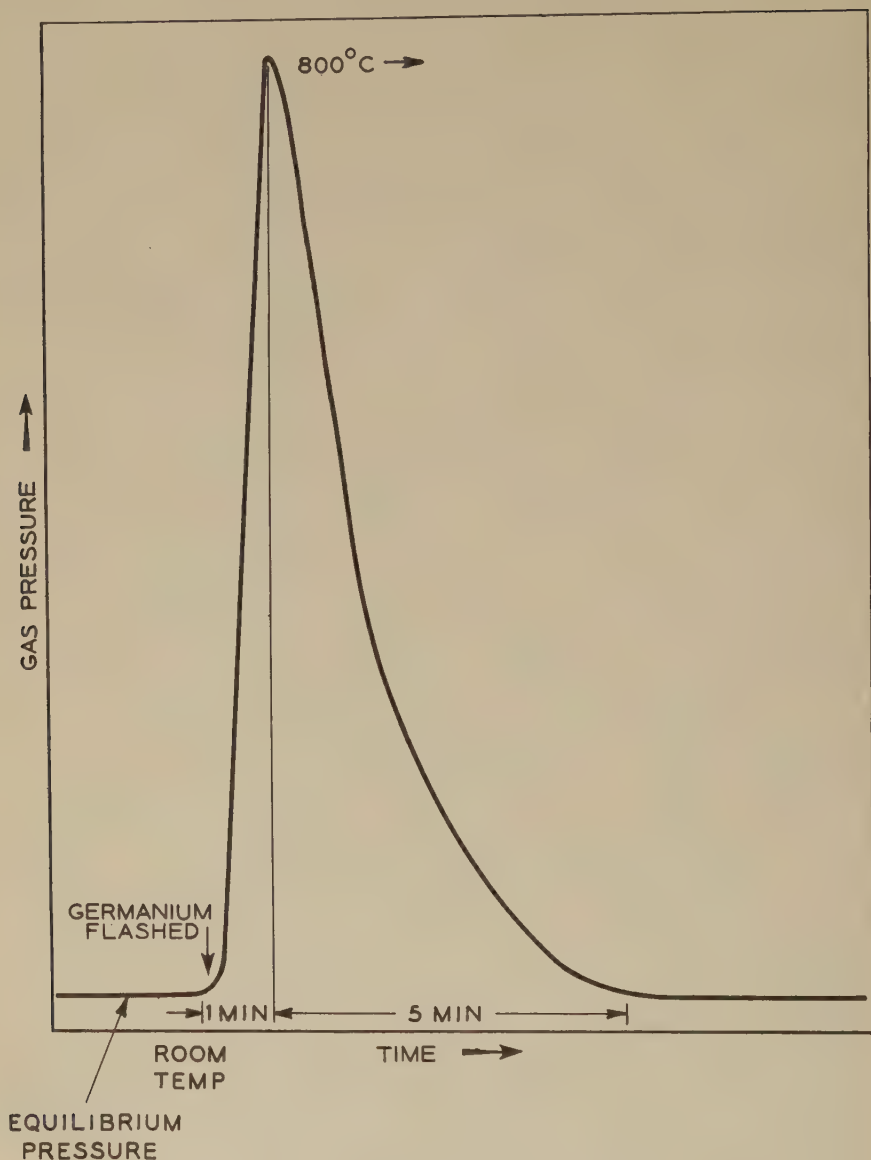


FIGURE 2. Typical pressure change on flashing.

pressure differs from one gas to another. For this reason, the instrument was calibrated for each gas used by establishing a known pressure (usually 10^{-3} mm.) on the high pressure side of the leak shown in FIGURE 1 and measuring the corresponding peak height obtained. The pressure drop across the leak is known or may be calculated for all the gases studied.

The volume of the system is known, so that the number of molecules of gas desorbed may be calculated from

$$M = \frac{KV\Delta p}{A}$$

where K is the number of molecules per liter for $p = 1$ mm., i.e., at 300°K , $K = 3.2 \times 10^{19}$; V is the volume of the system; Δp the change in pressure (i.e., peak height) and A is the area of the germanium in cm^2 .

The geometrical area of the CP-4 etched sample (filament 1) was 8 cm^2 and this figure was used for A . For the sandblasted sample (filament 2), however, the peak heights were all three times as large as for 1, so an area of 24 cm^2 has been used in the calculation of the number of molecules desorbed.

5.0. Results

5.1. Rare Gases and Nitrogen. Neon, argon, krypton, and nitrogen were introduced individually to the germanium filament as explained in the experimental procedure. In each case, no desorption was observed on flashing the germanium. It is therefore concluded that none of these gases are adsorbed on germanium at temperatures of 20°C . or higher.

5.2. Carbon Dioxide, Carbon Monoxide, and Hydrogen. The results obtained for the adsorption of these gases on germanium as a function of time are shown in FIGURE 3. In this case, the time shown represents the period for which the gas was allowed to remain in contact with the cold germanium before flashing. In each case, a pressure of 3×10^{-5} mm. Hg of gas over the filament was used. It will be seen that almost six minutes are required for the adsorption process to reach completion, which at the pressure used corresponds to a very low sticking probability ($\sim 10^{-3}$) for gas molecules striking the germanium. The relative amounts of carbon monoxide, carbon dioxide, and hydrogen desorbed are in the ratio of 8:2:1. Discussion of this variation will be left for a later paper.

No evidence was found for the conversion of any appreciable amounts of carbon dioxide adsorbed on the surface to carbon monoxide (cf.⁷).

Two curves for hydrogen, labeled respectively (from water) and (molecular) are given in FIGURE 3. The latter is the normal desorption of hydrogen which had been admitted as molecular hydrogen. It was found, however, that when water vapor was passed over hot germanium, hydrogen was formed. If the filament was then cooled and flashed, hydrogen was desorbed in quantities five times greater than when molecular hydrogen has been admitted. The obvious explanation for the difference is that, in one case, we are dealing with the adsorption of molecular hydrogen as such, while in the other atomic hydrogen is formed by the decomposition of the water on the filament; and this then chemisorbs as atoms. However, at the present time, we have no experimental proof of this assumption.

5.3. Oxygen. The reaction between oxygen and germanium has been studied by the method outlined in Section 3.3 because, when an attempt is made to flash adsorbed oxygen off the germanium, reaction occurs and the oxygen pressure in the system decreases. A known pressure of oxygen was allowed to flow contin-

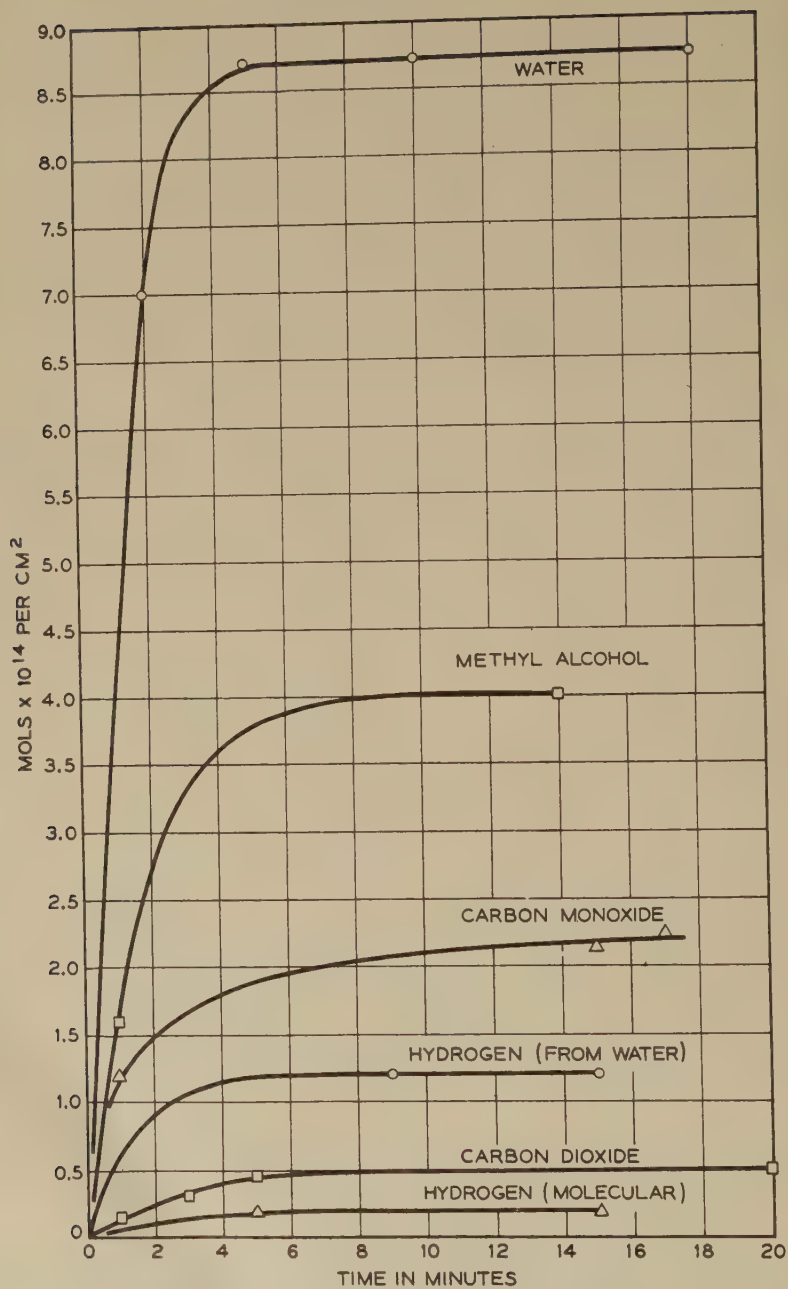


FIGURE 3. Desorption of gases from germanium.

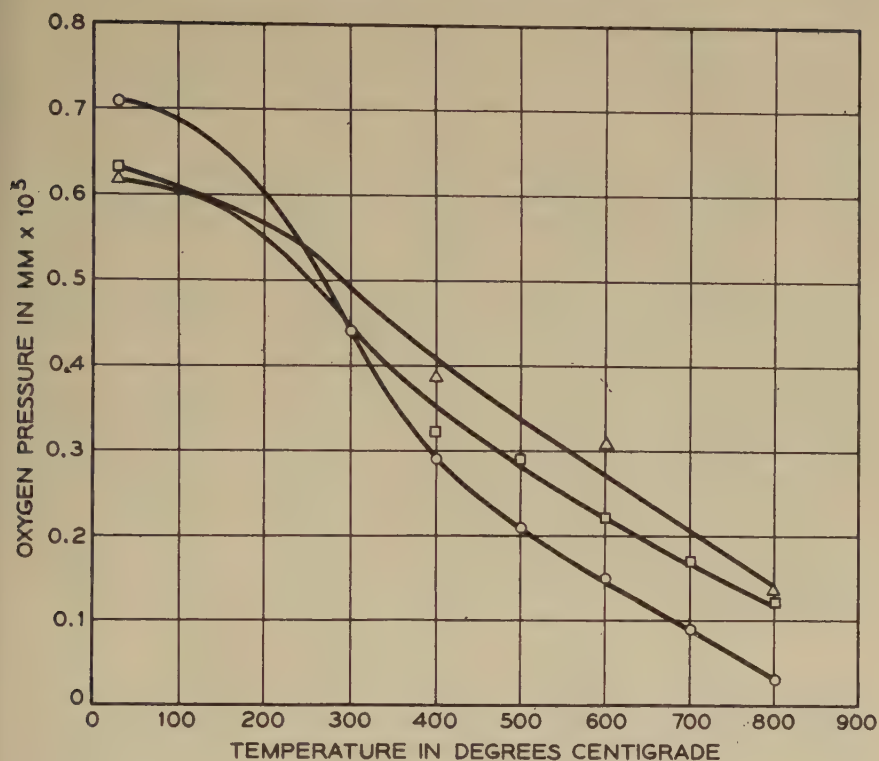


FIGURE 4. Effect of temperature on the reaction of oxygen and germanium.

uously over the filament, which was heated in steps, while the oxygen pressure was continuously recorded. The results are shown in FIGURE 4, where it will be seen that oxidation commences at relatively low temperatures $\sim 200^\circ \text{C}$. At a given temperature, the oxidation process appears to continue indefinitely so that either the film formed is porous or the oxide evaporates. With the apparatus used, the path between the filament vessel and the ion chamber was too long to measure concentrations of GeO or GeO_2 . It is hoped to continue this work later, using a more favorable geometry. With this reaction, as with water and methyl alcohol, the oxygen pressure measured and shown is probably considerably higher than the pressure in the filament vessel, as we have two types of pumping going on. The first type is the normal effect of the vacuum pumps, while the other type is caused by the germanium. The latter effect, attributable to reaction between oxygen and the germanium filament, differs from the pumping caused by adsorption in that it occurs at elevated temperatures where the measurements are being made. At these temperatures, no adsorption occurs, so that the filament pumps only when reaction occurs. As the germanium is not in the main line but rather on a side tube, higher pressures of oxygen may be present in the main line than in the side tube, owing to the constant admission of gas through the leak.

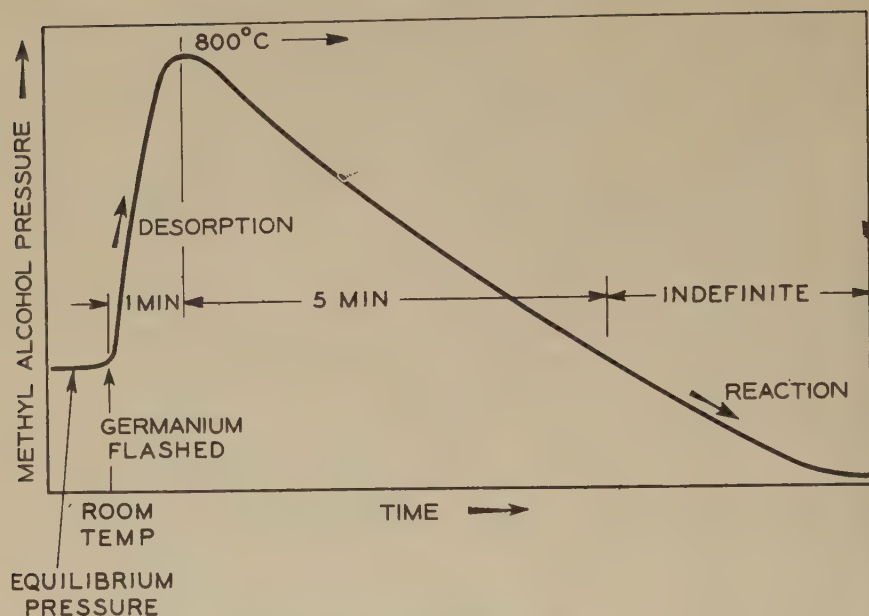


FIGURE 5. Desorption and decomposition of methyl alcohol vapor.

5.4. Methyl Alcohol. Methyl alcohol at a pressure of 3×10^{-5} mm. was passed over germanium for various time intervals, the system pumped and the filament flashed. The quantities of methyl alcohol desorbed are shown in FIGURE 3. The flash peak obtained had the shape shown in FIGURE 5 and the fact that the alcohol pressure with the filament hot falls below the value found with a cold filament indicates that decomposition of the methyl alcohol is occurring at the higher temperature. For this reason, background sweeps from mass 2 to mass 80 were carried out with methyl alcohol in the system and with the germanium both hot and cold. From these sweeps it was found that carbon monoxide and hydrogen were being produced at the germanium surface in appreciable quantities, probably as a result of the reaction:



The variation of the partial pressures of the methyl alcohol, carbon monoxide, and hydrogen were studied as a function of the temperature of the germanium. The results obtained are shown in FIGURE 6, and it will be seen that as the temperature is increased the decomposition of the alcohol is also increased. The pressures shown for the methyl alcohol are probably higher than those present in the filament vessel because the filament vessel is not in the path of the main pumping line.

5.5. Water. Values for the amount of water desorbed under given conditions are shown in FIGURE 3. The amount is dependent on the pressure used and, to some extent, on the time allowed for equilibration. It will be noted that water

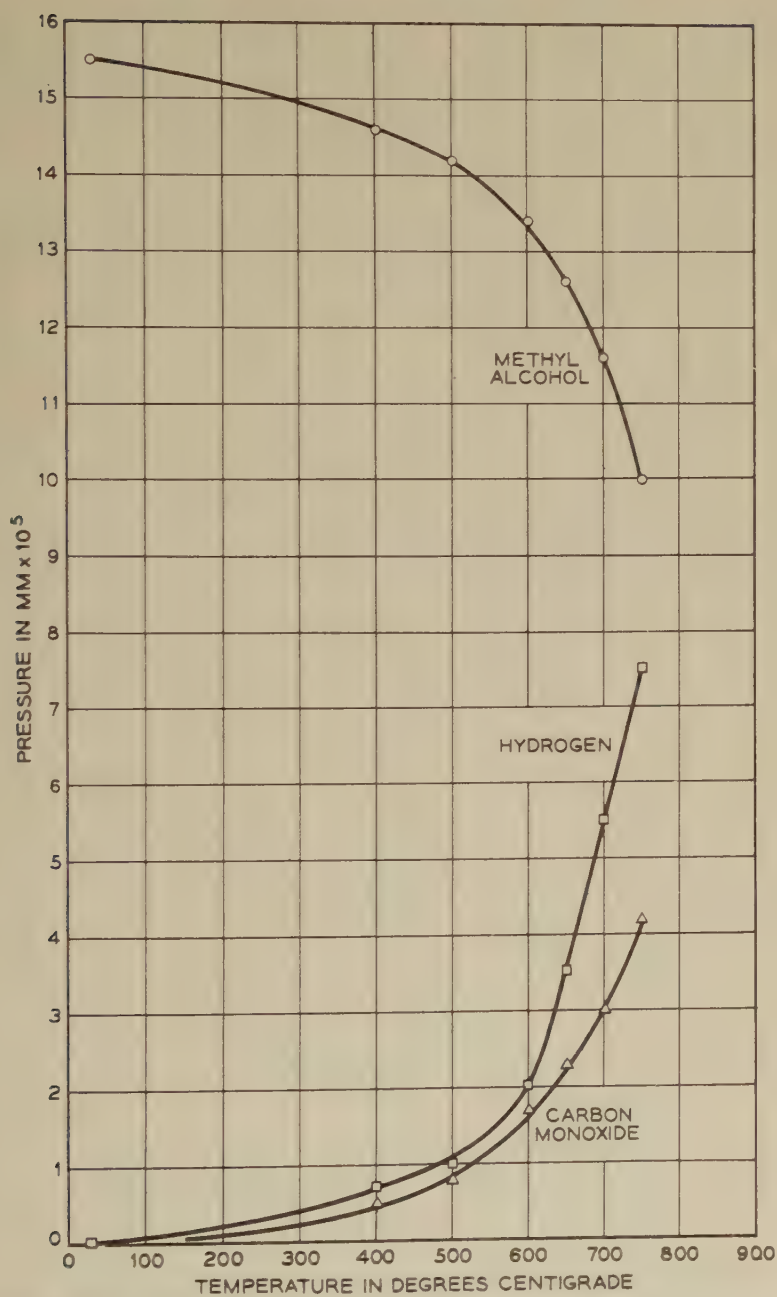


FIGURE 6. Effect of temperature on the decomposition of methyl alcohol vapor at a germanium surface.

adsorbs to a much greater extent than any gas or vapor studied, the amount desorbed being greater by a factor of five than that found for the majority of gases. The actual figures given for water may be at a lower equilibrium pressure than those listed because of adsorption on the glass walls of the apparatus. Because of this difficulty, one can say only that the pressure of water in equilibrium with the germanium was approximately 10^{-5} mm. of mercury.

The flash-off curves obtained for water had the shape shown in FIGURE 5 for methyl alcohol, *i.e.*, an initial peak followed by pumping below the original level. This result is attributable to reaction, and with water the only product observed was hydrogen, probably as a result of the following reactions:



These reactions were studied using the method outlined in Section 3.3; the results obtained are shown in FIGURE 7. Briefly, the method was as follows: water vapor at a known pressure was allowed to flow over the cold germanium and the mass spectrometer was focused on the hydrogen peak (mass-2). The germanium was then heated stepwise and a continuous record was made of the hydrogen partial pressure. This pressure increased as shown in the top half of FIGURE 7. The small peaks shown after each increase in temperature (except the first peak) are attributable to the desorption of hydrogen, which had ad-

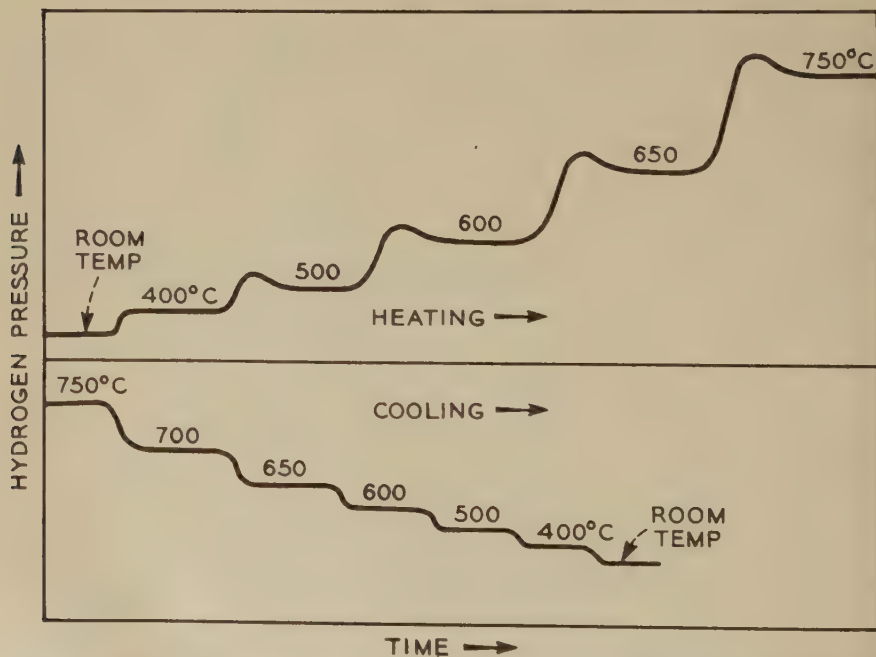


FIGURE 7. Variation of hydrogen pressure with temperature during the decomposition of water vapor at a germanium surface.

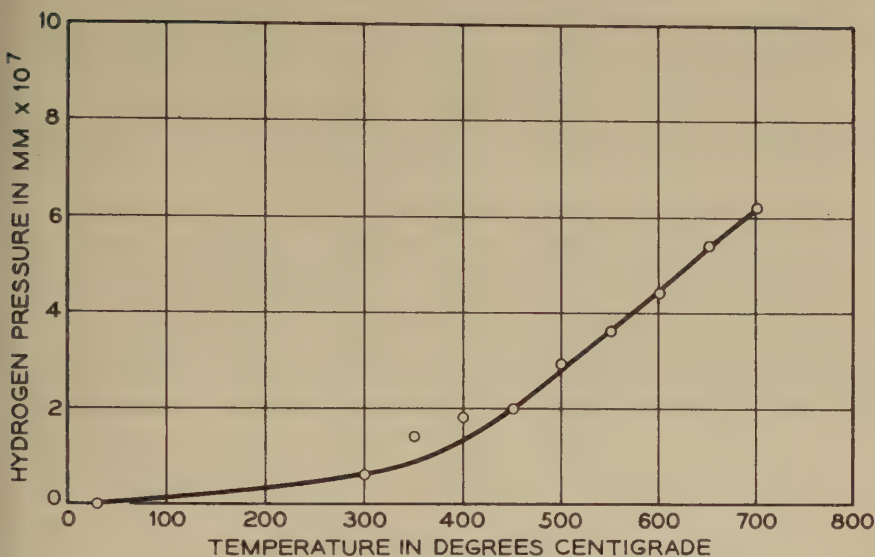


FIGURE 8. Reaction of water vapor with germanium.

sorbed on the germanium surface after decomposition of water. The absence of a peak in going from 20° C. to 400° C. indicated that no decomposition of water occurred at the lower temperature. Similarly, on cooling the germanium in a stream of water vapor, the hydrogen pressure decreased, as shown in the bottom half of FIGURE 7. The results obtained from heating or cooling were identical and are shown in FIGURE 8, where it will be seen that decomposition starts at temperatures as low as 250° C. and increases steadily as the temperature is raised. The ratio of hydrogen pressure to water pressure at 700° C. is of the order of 1:100 but again, because of adsorption of water on the glass, this ratio may be subject to error. At the present time, it is not possible to say definitely whether the monoxide or dioxide of germanium is produced although, at elevated temperatures, production of the former is the more likely.

6.0. Summary

The adsorption of neon, argon, krypton, nitrogen, hydrogen, carbon monoxide, carbon dioxide, oxygen, water, and methyl alcohol on germanium has been studied.

6.1. Neon, argon, krypton, and nitrogen showed no adsorption.

6.2. Carbon dioxide, carbon monoxide, and hydrogen were adsorbed on germanium at room temperature. Carbon monoxide adsorbed the most; carbon dioxide, less; and hydrogen, least.

6.3. Oxygen shows reaction in the form of oxidation of the germanium filament on flashing.

6.4. Methyl alcohol is adsorbed at room temperature and decomposes into carbon monoxide and hydrogen when the germanium filament is flashed.

6.5. Water is the most highly adsorbed of all the gases and vapors studied. Reaction occurs when the filament is flashed, giving hydrogen and a germanium oxide as its products.

6.6. The gases that adsorb on germanium are listed in order of decreasing adsorption: water, methyl alcohol, carbon monoxide, carbon dioxide, and hydrogen.

References

1. BRATTAIN, W. H. & J. BARDEEN. 1953. Bell System Tech. J. **32**: 1.
2. APKER, L. 1948. Ind. Eng. Chem. **40**: 5, 846.
3. MOLNAR, J. P. & C. D. HARTMAN. 1950. M.I.T. Physical Electronics Conference. April 1950.
4. BECKER, J. A. & C. D. HARTMAN. American Physical Society Meeting, June 15, 1951. Schenectady, New York.
5. PFANN, W. G. & K. M. OLSEN. 1954. Phys. Rev. **89**: 322.
6. CONWELL, E. M. 1952. Proc. I.R.E. **40**: 1327.
7. KWAN, T. & Y. FUJITA. 1953. Nature. **171**: 705.

ELECTRICAL CONDUCTIVITY OF GERMANIUM SURFACES

By Edward N. Clarke

Sylvania Electric Products Inc., Bayside, N. Y.

Introduction

The bulk properties of germanium have been extensively studied.¹ It is well known that chemical impurities and lattice defects introduce new energy levels within the solid. The measurements of resistivity and Hall coefficient determine mobility and carrier concentration, and have provided a valuable method for determining the location of some of the energy levels within the solid. It is to be expected that chemical impurities and lattice defects on the surface (of course, the surface is in itself a major defect) might introduce new energy levels. It is important to determine the detailed nature of these surface levels because they play a major role in such things as carrier trapping, excess carrier recombination, surface potentials, noise, and surface conductivity. All of these properties are important in the field of electronics involving rectification and transistor action.

Earlier work,^{2, 3, 4, 5} which dealt principally with contact potentials, suggested that surface levels did exist on the semiconductors silicon and germanium. This work did not provide detailed knowledge of the surface energy level distribution, although the later work of Brattain and Bardeen⁶ includes a more detailed model of surface energy levels.

Our own work is directed toward obtaining detailed information concerning surface energy levels. For example, we should like to determine the levels corresponding to the clean, perfect germanium surface and the levels corresponding to a surface covered with a specific impurity such as oxygen (or oxide layer) or the surface with certain defects. Several different approaches are possible, and the present paper reports on one approach that is being found useful. If surface conductivity and Hall coefficient can be made large enough to measure, then we ought to be able to determine surface carrier concentration and carrier mobility. Some of the surface energy level structure will be determined from the temperature dependence of conductivity or carrier concentration. We have found that the surface conductivity on germanium can be increased by mechanically disturbing the surface⁷ and also by introducing oxygen to an etched germanium surface partially cleaned by heating in a high vacuum.⁸

Results

Mechanically induced conductivity. If the germanium surface is mechanically disturbed, as by grinding, polishing, or sandblasting, we have, in effect, a surface layer in parallel with the bulk crystal. A comparison of resistance before and after sandblasting will allow determination of the surface layer resistance. FIGURE 1 shows resistivity plotted on a log scale as a function of the reciprocal of absolute temperature in the temperature range 385° K. down to 78° K. By effective resistivity, we mean that of the composite crystal, surface plus volume. The upper curve represents the crystal with etched surface and the lower curve

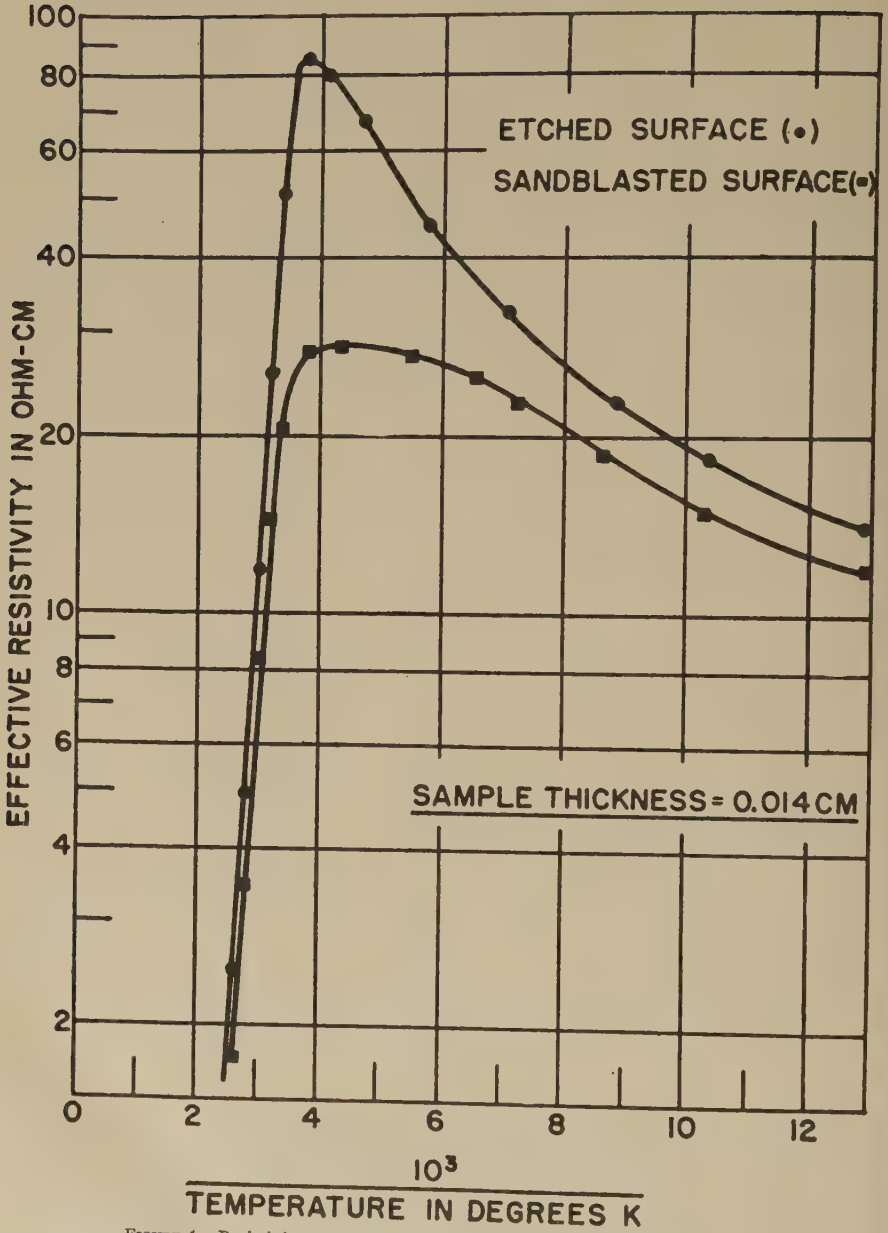


FIGURE 1. Resistivity of a thin single crystal before and after sandblasting.

that with one surface sandblasted with 3 micron silica particles at about 80 pounds per square inch pressure. The effective room temperature resistivity is decreased from 47 ohm-cm. to 20 ohm-cm. The change in effective resistivity produced by going from one surface to the other is amplified by using a thin crystal and the thickness of this particular single crystal is 0.014 cm. This plot shows a decrease in effective resistivity by as much as a factor of 3. It is to be noted that, even at the highest temperature, the resistivity with a sandblasted surface is lower than that with an etched surface, implying that there exists a large concentration of extrinsic carriers in the surface layer. Now, if we assume that the surface layer is homogeneous, we can get an average surface layer resistivity, having measured surface resistance and surface layer thickness. The surface layer thickness is estimated by etching the surface slowly and watching the effective room temperature resistivity return to its original value. The etchant has been calibrated and returns the effective resistivity to its original value after about 45 minutes. The surface layer thickness determined in this way is about 0.7×10^{-4} cm.

The disturbed surface layer resistivity varies with temperature, as shown in FIGURE 2. Note that the resistivity at room temperature is about 0.2 ohm-cm. Of course, the absolute value of the resistivity depends upon the detailed procedure used in disturbing the surface. There are three distinct regions. In the range 345° K. to 385° K., the slope corresponds to an ionization energy of 0.4 electron volt and, in the range 143° K. to 300° K., to 0.022 electron volt. From 83° K. to 125° K. there is a small negative slope. Of course, this neglects the temperature variation of mobility. The value 0.4 ev. must be considered a preliminary one, because it is based upon only two temperatures for each surface and is thus only roughly reproducible with a given sandblasting procedure.

Thermoelectric measurements with a hot point probe have shown that the mechanically disturbed surface layer is *p*-type, at least in the region near room temperature.⁹ If the hot point probe technique is used with care, we have found that the disturbed surface is *p*-type to temperatures as high as those covered by the present experiment on conductivity. Brattain and Bardeen⁶ have found that the contact potential between germanium and a reference platinum electrode is decreased (or the germanium work function increased) with sandblasting of the germanium surface. This discovery is, of course, in agreement with the finding that a disturbed surface is *p*-type. It is therefore tentatively assumed that the apparent ionization energies 0.4 and 0.022 electron volts represent two acceptor levels.

A number of workers, including Navon, Bray, and Fan¹⁰ have shown that the rate of recombination of electrons and holes at the surface of germanium is increased by a large factor if the germanium surface is mechanically disturbed. We accordingly suggest further that the high-lying level represents the centers responsible for the recombination of electrons and holes at this surface.

If we assume that mobility in the surface layer is about the same as mobility in the bulk, then carrier concentration includes the range 10^{15} to 10^{17} carriers per cm³. If the mobility is actually lower, then the carrier concentration is correspondingly higher. Of course, a measurement of surface Hall coefficient in

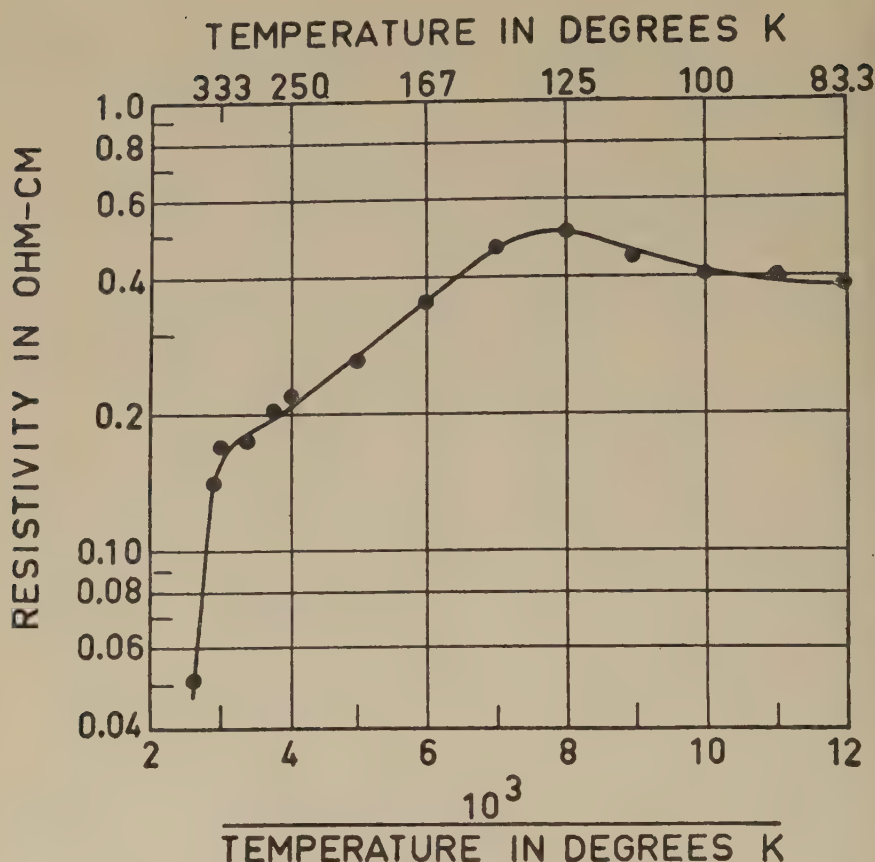


FIGURE 2. Resistivity of the mechanically disturbed surface layer.

order to determine surface carrier concentration would be a most valuable addition to the present work. We have attempted to measure surface Hall coefficient, and FIGURE 3 shows the effective Hall coefficient, *i.e.*, that of the composite crystal, surface plus volume. This is the same sample with thickness 0.014 cm. It is seen that sandblasting reduces the effective room temperature Hall coefficient by a factor of about 5 (the sign remains unchanged). The reduction is as large as a factor of 8. Curves of this type are still being analyzed in order to determine the Hall coefficient in the disturbed layer, but we have thus far not obtained consistent results.

We have tried sandblasting with larger particles, sandblasting with a different material such as silicon carbide, lapping with fine grit, with polishing paper, and even polishing with rouge and water on felt. Surface conductivity is increased with all these treatments, but the absolute value of surface resistance may vary by as much as a factor of 10 depending upon the detailed method used.

The disturbed surface can be annealed by heating at high temperatures, and

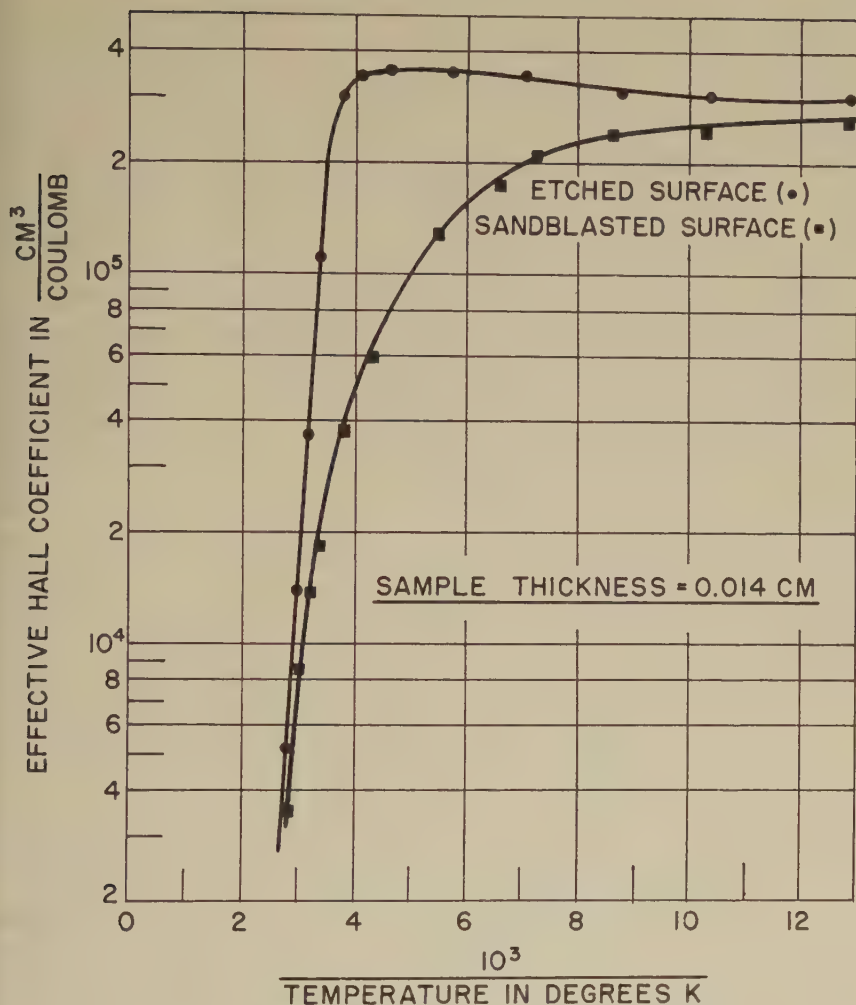


FIGURE 3. Hall coefficient of a thin single crystal before and after sandblasting.

FIGURE 4 shows the surface layer resistivity of a different sample increasing with annealing time at 488°C . in helium. A surface layer thickness of 10^{-4} cm. is assumed. The experimental accuracy is such that it can be said only that the surface layer resistivity returns to some value greater than 10 ohm-cm. after 21 minutes. The curve is particularly interesting in that it shows a faster annealing rate becoming operative only after annealing has proceeded for 10 minutes. There are several possible explanations for the shape of this curve. For example, a nucleation process might be active. On the other hand, we are also reminded of the plastic deformation of germanium which requires an "incubation" time before the external stress becomes effective.¹¹ The mechanically

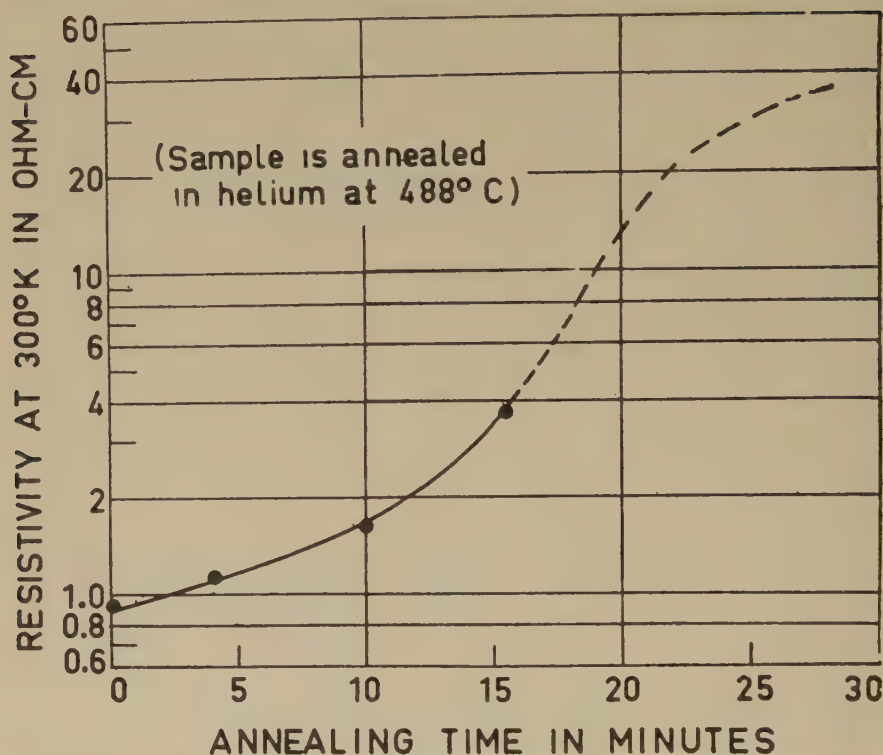


FIGURE 4. Resistivity of the mechanically disturbed surface layer as a function of annealing time.

disturbed germanium surface is under large stress (compression) as shown by the observation that a very thin (10^{-3} cm.) crystal is bent about 10° by the disturbed surface layer. The stress provided by the base material may thus play an important role in the annealing process. Other effects, however, such as the reaction of germanium with surface impurities might alter the shape of the annealing curve. The annealing curves are not always of this simple shape. Others are of such shape as to suggest that when a larger concentration of centers is present, several different kinds of disturbance centers with corresponding different energies are also present.

It is of interest to compare our results with those obtained after damage of the germanium crystal caused by fast neutrons.¹² Acceptor levels are introduced and the Hall data suggest apparent ionization energies of 0.14 and 0.31 ev. although there may really be a wide distribution of ionization energies. These acceptor levels are attributed to vacancies. The observation that a pair of levels is introduced might suggest that some of the centers introduced by sandblasting or polishing are similar to those produced by fast neutrons (interstitials and vacancies).

Our work on mechanically disturbed germanium surfaces is being continued in order to improve the measurements at high temperature, and also in order

to obtain the Hall coefficient for the disturbed surface layer and hence carrier concentration and carrier mobility; also, work is continuing on the annealing of these disturbed surface layers in order to suggest a possible mechanical structure for the centers responsible for the surface conductivity.

Oxygen induced conductivity. The conductivity of the etched germanium surface is being studied by heating the germanium in a high vacuum system and then introducing various gases. Mott¹³ has suggested that oxygen may introduce acceptor energy levels on a germanium surface, and we should like to look into the details of this possible process.

A very thin (0.004 cm.) high resistivity (40 ohm-cm. at room temperature) n-type single crystal was assembled for electrical measurements in a high vacuum system. The surface was prepared by etching with an etchant of HF, HNO₃, and H₂O, and then rinsing with H₂O. The residual gas pressure was reduced to about 5×10^{-8} mm. Hg after pumping for 20 hours and after heating the experimental tube containing the germanium to about 450° C. for one half-hour. This treatment decreased the germanium conductance by a factor

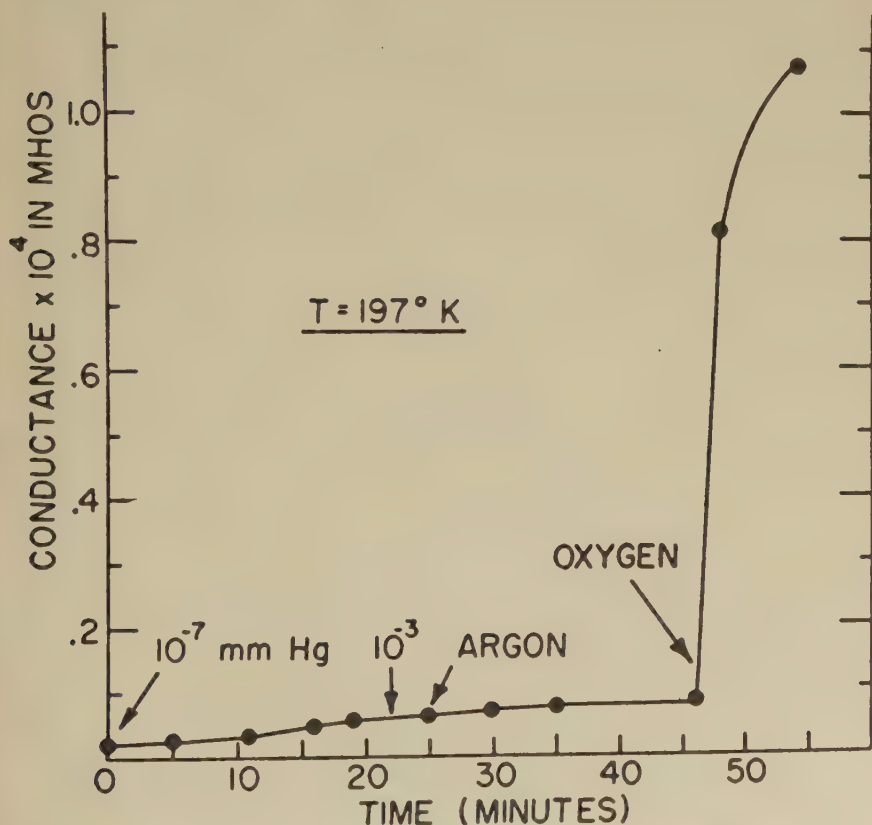


FIGURE 5. Dependence of Ge conductance upon surrounding atmosphere. The diffusion pump was shut off at time $t = 0$; argon was admitted at $t = 25$ minutes and removed after one minute; oxygen was admitted at $t = 47$ minutes and removed after one minute.

of 2, representing a decrease in conductance of 4×10^{-6} mho. If one assumes a surface carrier mobility of $1000 \text{ cm}^2/\text{volt sec.}$, this diminution represents a decrease of 4×10^{10} carriers per cm^2 of surface. A different carrier concentration will, of course, correspond to a different mobility.

The germanium was heated to near its melting point for a short period of time (seconds), with a resulting increase in residual gas pressure, and then quenched rapidly by radiation to the surroundings at room temperature. This heating and quenching in a high vacuum may have produced a surface cleaner than the original etched surface, but they also changed the bulk characteristics. In particular, much of the crystal was converted to *p*-type. After this heating and quenching, gases at several different pressures were admitted slowly through the vacuum system. FIGURE 5 shows a typical plot of germanium conductance at 197° K. , varying as the residual gas pressure varies from 10^{-7} mm. Hg. to 10^{-3} mm. Hg. by shutting off the diffusion pump. Dry argon admitted to the germanium at 0.3 mm. Hg. for 1 minute has little effect upon the conductance. Dry oxygen at 0.3 mm. Hg. for 1 minute increases the conductance by a factor of 10. The largest surface conductance at 197° K. obtained during similar observations is 8×10^{-4} mho, representing about 10^{13} carriers per cm^2 of surface, if a mobility of $1000 \text{ cm}^2/\text{volt sec.}$ is assumed. This would suggest that

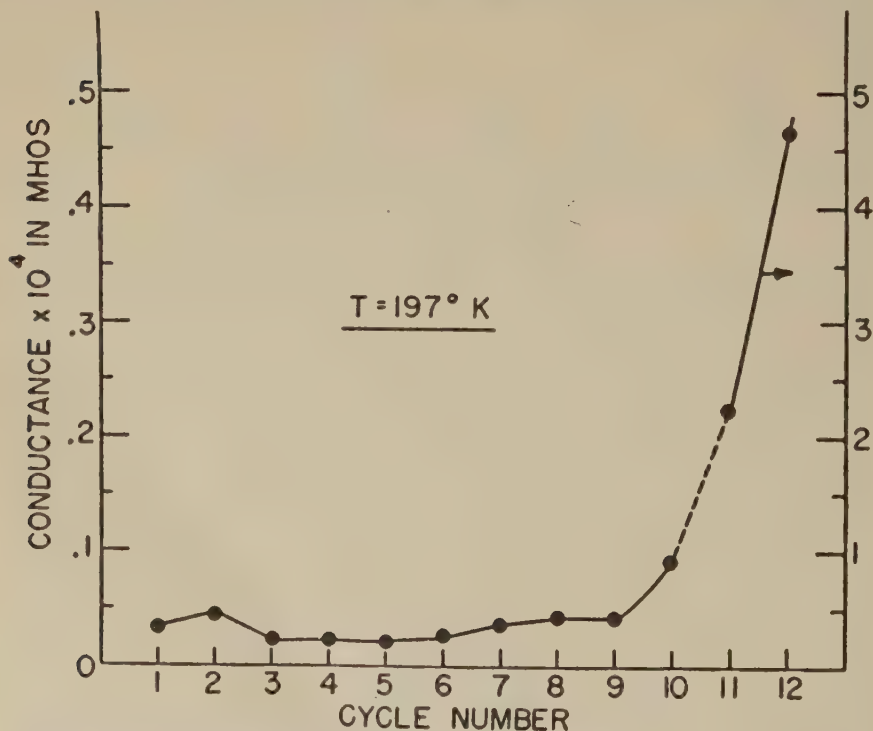


FIGURE 6. Conductance of Ge after heating in high vacuum. Each cycle corresponds to a maximum and minimum conductance. Cycles 11 and 12 correspond to right-hand ordinate.

the Fermi level is about 0.08 ev. above the valence band at the surface (if we have hole conductivity).

Initially, it appeared possible to remove most of the adsorbed gas (as measured by electrical conductance) by simply heating the experimental tube containing the germanium in a high vacuum at about 450° C. for one half-hour. However, FIGURE 6 shows that, as the number of cycles increased (each cycle corresponding to a maximum conductance with gas adsorption, and then a minimum conductance after heating in high vacuum), the minimum conductance became larger. By the twelfth cycle, heating the tube at 450° C. had a much smaller effect in reducing the surface conductance. This might be described in terms of a thin layer (oxide, for example) being built up on the surface. Some of the oxygen might leave the surface by diffusing into the germanium rather than by entering the vacuum system. Further work will include measurements to determine surface carrier type and surface mobility of carriers as well as surface energy level structure introduced by oxygen.

After this paper was submitted for publication, the author obtained additional evidence for the important role of oxygen on the germanium surface. If oxygen is added to a surface having *n*-type conductivity rather than *p*-type conductivity, then the conductivity will decrease rather than increase with electrons from the conduction band becoming trapped in the oxygen levels. In addition, there is some evidence that oxygen plays an additional role. Oxygen may introduce donor levels when in solution in the germanium as compared with its role as an acceptor level or electron trap when adsorbed on the surface. This means that oxygen may be able to account for both donor and acceptor type surface levels on germanium, with the donor levels found just below the germanium surface and the acceptor levels on the surface (possibly with a very thin oxide layer separating the two groups of levels). This latest work is covered in Clarke, E. N. 1954. *Phys. Rev.* **95**: 284.

References

1. CONWELL, E. M. 1952. *Proc. I. R. E.* **40**: 1327.
2. MEYERHOFF, W. E. 1947. *Phys. Rev.* **71**: 727.
3. SHOCKLEY, W. & G. L. PEARSON. 1948. *Phys. Rev.* **74**: 232.
4. BARDEEN, J. 1947. *Phys. Rev.* **71**: 717.
5. BRATTAIN, W. H. & W. SHOCKLEY. 1947. *Phys. Rev.* **72**: 345.
6. BRATTAIN, W. H. & J. BARDEEN. 1953. *Bell System Tech. J.* **32**: 1.
7. CLARKE, E. N. & R. L. HOPKINS. 1953. *Phys. Rev.* **91**: 1566.
8. CLARKE, E. N. 1953. *Phys. Rev.* **91**: 756.
9. GRANVILLE, J. W. & C. A. HOGARTH. 1951. *Proc. Phys. Soc. London, England.* **B64**: 488.
10. NAVON, D., R. BRAY, & H. Y. FAN. 1952. *Proc. I. R. E.* **40**: 1342.
11. GALLAGHER, C. J. 1952. *Phys. Rev.* **88**: 721.
12. CLELAND, J. W., J. H. CRAWFORD, JR., K. LARK-HOROVITZ, J. C. PIGG, & F. W. YOUNG. JR. 1951. *Phys. Rev.* **83**: 312.
13. MOTT, N. F. 1949. *Proc. Inst. Elec. Engrs. London, England.* **96**: 253.

EXPLORATION OF METAL SURFACES WITH FINE WIRES

By R. H. Savage and D. G. Flom

Research Laboratory, General Electric Company, Schenectady, N. Y.

When fine* wires of platinum or tungsten are touched together under a microscope, they may show unexpectedly strong cohesion, especially if cleaned previously as by heating to a bright color. The cohesion, however, may show this peculiarity: that the wires are attracted sufficiently to remain in contact while one is being pulled normally away, and simultaneously one wire may be slid laterally upon the second without destroying this attraction.† As an explanation of this effect, the capillary attraction of liquid-film contaminants is unlikely in view of the cleaning precautions, and welding (as known to occur between clean metals) is not indicated where lateral sliding can occur.

Cohesion of this kind, concurrent with a kind of lubrication, is not observed with large wires, and it is described here as an illustration of certain mechanical effects which occur at surfaces on a micro-scale which are normally masked by the mass or stiffness of grosser objects.

Further microscopic observations of this kind with electrolytically pointed wires suggested to one of us (RHS) that the more fragile films on metal surfaces may have important properties undetected by ordinary exploring probes which may puncture them mechanically. From this viewpoint, it seemed essential to measure and control sensitively the normal force of the fine wire against other surfaces. We decided to make measurements by employing the elastic properties of the wires when shaped into loops.

Preliminary experiments indicated that a 2-mil wire, forming a loop of about 1-cm. diameter, served as a sensitive contact probe. Platinum was chosen initially to avoid high resistance films in the measuring circuit, and gold was used alternatively. The loop was not quite closed at the top. As shown in FIGURE 1, the measuring current was passed down one side of the loop while the potential drop at the contact was measured down the other side with a potentiometer. This circuitry was necessary to avoid the confusion of supplementary resistances along the wire as far as the junction.

The force of the loop against the specimen was indicated from the elastic compression measured with a cathetometer, microscope, or micrometer. These devices were tried interchangeably and each one proved satisfactory. The force of the deformed loop at any given amount of compression had been determined by calibrating the loop directly against the pan of an analytical balance. A typical calibration curve is shown in FIGURE 3. The sensitivity can be lowered easily into the microgram range.

We had intended to study with this apparatus the properties of various films of known thickness and composition, but, at this stage, two practical problems appeared which were concerned with possible stray films of unknown origin on faulty metal contacts taken from service. This fortunate circumstance invited immediate trial of the exploratory possibilities of wire-loop probes.

* Diameter 2-mils or less.

† We have not found a reference to this peculiar effect.

The first problem involved a button of a platinum-osmium alloy making intermittent contact with one of "carboly." The latter was claimed to have been coated with platinum-osmium during contact, and was intended to be relatively film-free; but, in service, this contact device had shown apparent mechanical closing in many instances without completing the corresponding electrical contact. Ordinary coarse wire probes applied to the surfaces in question showed negligible film resistance. However, a platinum loop, delicately applied as shown in FIGURE 2, showed small areas having resistance as high as 10 megohms. With the aid of charged condensers, it was found that these high resistances would withstand 30 volts without breakdown, and thus would account easily for the observed contact failures.

The second problem involved the surfaces of several large silver contacts. Some of these contacts had been buffed and others had been machined. Only the latter had proven to be satisfactory in a low voltage circuit.

Upon exploring these surfaces with wire loops, we found that the buffed surfaces showed very many high resistance areas which were not changed by mild cleaning. On the other hand, the machined surfaces showed consistently lower surface resistance against the loops, and the areas which were higher than normal in resistance were easily lowered by mild cleaning with organic solvents. Contamination from handling proved to be the source of the higher resistance films in this case. We found that transparent fingerprints could be detected on otherwise clean silver surfaces by this technique.

In these problems, it became important to find some basis for comparing film strengths. Experiment then showed that a single high-resistance contact surface under a loop could be broken down either by increasing the normal force at constant voltage or by increasing the voltage at constant force. Both methods yield useful information, but a third type of comparison is also worth while. This type consists in comparing local contact resistances over a series of many spots, applying constant voltage and force to each one in succession.

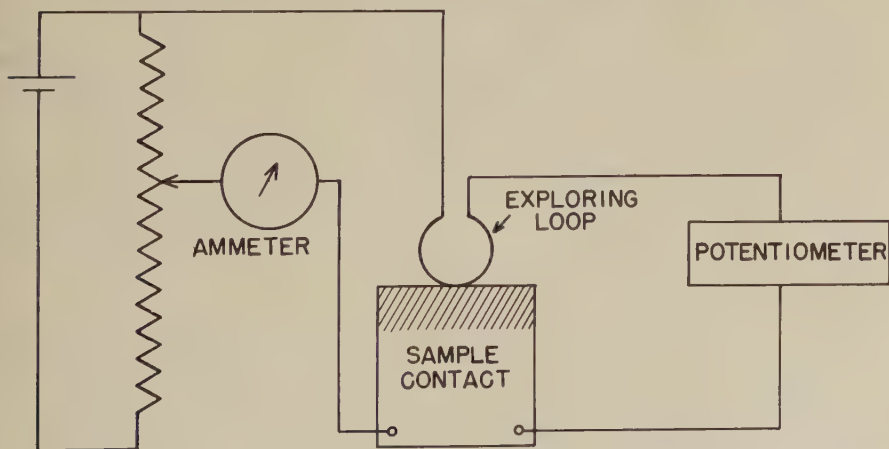


FIGURE 1. Circuit for measuring contact resistance with wire-loop probe.



FIGURE 2. Exploration of contact surface with wire-loop probe.

The basis for this latter comparison lies in the theoretical notion that the area of actual contact, as distinguished from the apparent area, is directly proportional to the normal force upon the loop. This assumption follows from Hertz's equations¹ for areas of contact between adjoining elastic bodies, which include also the radii of curvature, Poisson's ratio, and Young's elastic moduli for the two solid bodies. Calculations of this kind constitute in themselves a separate subject, out of line with our present discussion, but the influence of their underlying principles upon our experimental work should be easily seen and a few points at least noted here.

The calculation of actual contact areas enables one to determine the tensile strength of the contact-adhesion referred to in our introductory paragraph and thus throws light on the nature of the attractive forces. It also serves to indicate the magnitude of the contact pressure and the influence of this pressure on the physical properties of the material in the junction. The mechanical properties of thin films may thus be examined.

Even though a loop may be pressed against a surface with a low total force, it does not necessarily follow that the pressure will be low. In fact, the calculation of the Hertz contact area (under a known applied force and for a given geometry of loop) has revealed that pressures corresponding to the macro-hardness of the contact materials may be produced. This is rather to be expected from Hertz's assumptions, which utilize the elastic constants to the limit before plastic flow begins.

The ease with which such calculations can be made depends considerably upon the geometry of the contact members. For a circular loop, the force at the contact interface will be applied at a single point only upon *initial* contact. Further compression will result in distribution of the force along a defi-

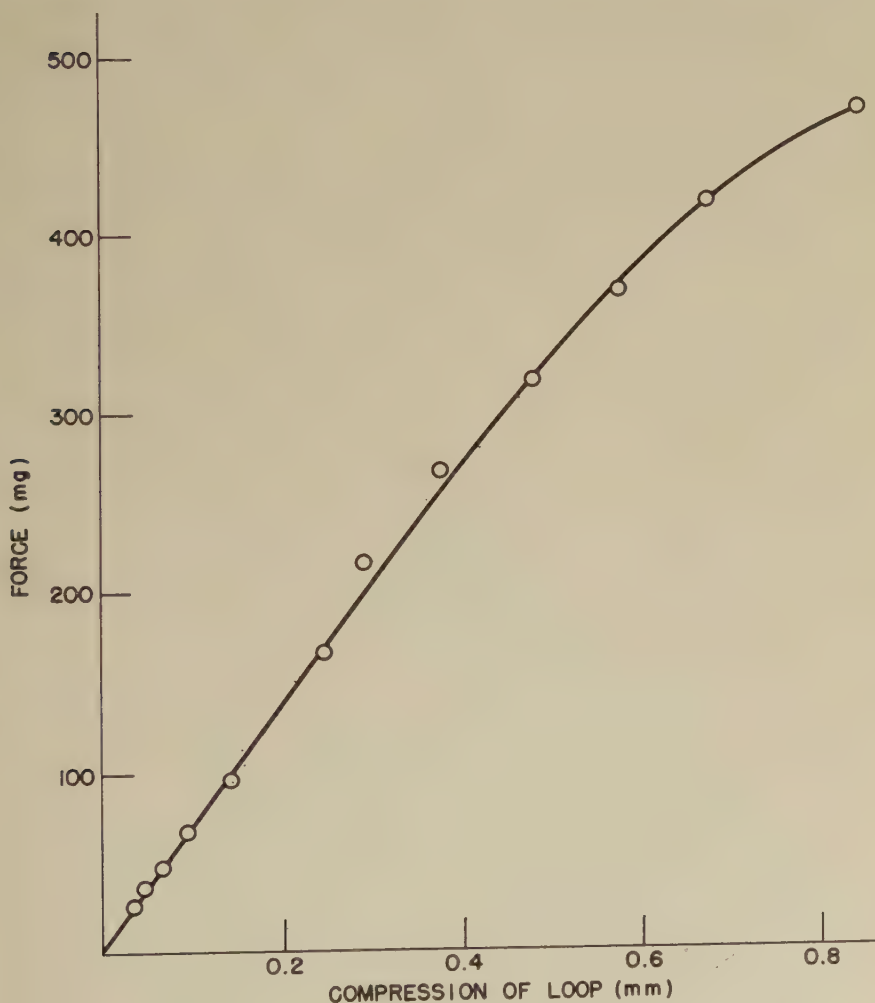


FIGURE 3. Calibration curve for 2-mil platinum wire loop.

nite length of the wire, and a complication arises from the changing radius. But if a loop is used in which a small round protrusion has been shaped at the contact zone, distortion of this minor curve can be avoided and the calculation is simplified.

At best, a calculation of area from the Hertz equations should be considered as only suggestive, especially when the wire diameter is made very small. The results with 2-mil wires are not in agreement with those based on electrical resistance measurements entered into the well-known equations of Maxwell.

In conclusion, this paper is intended to show that a very sensitive method of exploring small regions ($<10^{-8}$ cm.²) on conducting surfaces originates in the use of fine curved-wire probes. From the practical problems already cited, in which the method was developed, we have extended the investigation to include metal surfaces and fine wires cleaned in vacuum and subjected to known atmospheres. For example, surface inhomogeneities on gold surfaces, fired to the melting point, are easily found in this way. More of this work will be described in a later paper.

Reference

1. F. P. BOWDEN & D. TABOR. 1950. Friction and Lubrication of Solids. Oxford Univ. Press. London, England. Ch. 1. Also H. M. WESTERGAARD. 1952. Theory of Elasticity and Plasticity. Wiley. New York, N. Y. Ch. 6.

SURFACE PROPERTIES OF GERMANIUM AND SILICON

By W. H. Brattain and C. G. B. Garrett
Bell Telephone Laboratories, Murray Hill, N. J.

The early work on the thermionic emission from clean tungsten surfaces, and the dependence of the emission on adsorbed ions such as cesium, thorium, oxygen, *etc.*, has served as a landmark and reference point for the understanding of surface phenomena.¹ At the present time, a clean surface of a single face of a tungsten crystal in vacuum is the best understood surface, and forms a convenient starting point from which to proceed in the understanding of more complicated surfaces. Such a procedure, however, becomes less and less profitable, the further a given surface departs from this ideal, *i.e.*, the dirtier the surface. It has been clear for some time that the surface of a semiconductor exposed to a gas or liquid environment represents a very considerable departure from the ideal of a clean tungsten crystal in high vacuum. For such a system, fortunately, it turns out that the electrical properties may be more easily understood by comparison with an entirely different kind of interface.

If one considers the various possible interfaces, namely, all the permutations, two at a time, of solid, liquid, gas, and vacuum, the furthest extreme from the solid-vacuum would be the solid-solid interface. Recent progress in the understanding of semiconductors has supplied such a reference point. The interface between *n* and *p*-type germanium (a *p-n* junction) that can occur in a single crystal of germanium, is as simple and well understood now as the tungsten-vacuum interface and serves as another reference point or ideal at the opposite pole, so to speak.² This conception is an advance in the sense that one now has ideal interfaces at both extremes and one can approach the less understood middle from either end.

As might be expected, the *p-n* junction is a particularly appropriate reference-point for semiconductor surfaces, especially for those of germanium and silicon.

The germanium crystal has the diamond structure, the atoms being held together by covalent bonds. Each germanium atom shares one of its valence electrons with each of its four nearest neighbors. In the perfect crystal at the absolute zero, all electrons are in the valence bonds, so that no electrons are left over for electrical conduction. The crystal is a good insulator. However, the energy E_g required to free an electron from the valence bond is finite. Excitation of an electron has, therefore, a finite probability at nonzero temperature. It turns out that both the excited electron and the hole left behind are free to drift in an electric field and thus to conduct electricity. This situation is quite analogous to that of ions in dilute solutions. Pure water supercooled to absolute zero would be a good insulator, but at finite temperatures, a certain fraction of the H_2O molecules will be dissociated into H^+ and OH^- , both of which are mobile ions which can contribute to the conduction of electricity. As a matter of fact, the excitation energy is about the same for germanium as for water, 0.7 and 0.5 electron volts respectively or, in other words, about 10 Kcal./mole. In each case, therefore, thermal energy maintains a certain number of

positive mobile carriers and an equal number of negative mobile carriers. One important difference should be noted. For germanium, the effect of illuminating the material with light of short enough wave-length (1.5×10^{-4} cm. or less) is to create an excess number of hole-electron pairs over the equilibrium value. These excess carriers have a short but finite lifetime (of the order of 10^{-3} to 10^{-6} secs.). They therefore can add appreciably to the electrical conductivity of the material. No such effect has been observed for water, presumably because the optically excited ion-pair has very little chance of escaping its surroundings and contributing to the electrical conductivity before recombination takes place.

For holes and electrons in germanium, just as for positive and negative ions in water, flow of mobile carriers takes place by diffusion owing to concentration gradients and by field flow. Except for the fact that the hole and the electron have much smaller effective masses and much higher mobilities, there is a very close similarity between the two systems.

At any nonzero temperature, the equilibrium concentrations p and n of holes and electrons are governed by the mass action law, just as in the electrolyte case.³ The product of concentrations is given in terms of the dissociation or excitation energy by

$$np = K \exp (-E_g/kT). \quad (1)$$

Now another similarity between the two systems appears. When an acid or an alkali is added to water, the equilibrium given by the mass action law is pushed very far to one side: either the (H^+) is much larger than the (OH^-) or vice versa, but the product of (H^+) and (OH^-) must be a constant for equilibrium at any given temperature. For example, adding HCl to water will result in a high concentration of H^+ and a low concentration of OH^- . Similar considerations hold for germanium. Atoms of elements from the fifth column in the periodic table, such as antimony, dissolve in the germanium in such a way that they go into the lattice substitutionally. Such atoms have 5 valence electrons, instead of 4, and the energy necessary to free the fifth electron is very small. Thus a solution of antimony in germanium behaves like a strong electrolyte in the sense that the added material is almost completely ionized at room temperature. Consequently, small amounts of antimony will increase the electron concentration and, therefore from EQUATION 1 proportionately decrease the hole concentration. Such a solution is referred to as n -type germanium. If, on the other hand, an element from the third column, such as gallium, is dissolved in the germanium, the effect is to increase the number of holes and decrease the number of electrons, giving p -type germanium. Solutions of equal numbers of atoms from columns 3 and 5 "compensate" each other, thus maintaining the equality of n and p as for pure germanium, just as the act of dissolving a neutral salt like KCl in water does not disturb the equality of (H^+) and (OH^-) . One difference between the two cases is that, in water, the *other* ions added (*i.e.* other than the H^+ and OH^-) are also free to move and contribute to the conductivity, whereas for germanium the ionized impurity atoms (such as Sb or Ga) are not, since at ordinary temperatures they are frozen in the lattice. The

fifth and third column impurities are called donor and acceptor atoms respectively, and their concentrations are generally represented by N_D and N_A .

In view of these similarities, it is not surprising that the mathematical formalism which has been developed for handling problems connected with the electrical properties of semiconductors bears a considerable resemblance to that developed long ago for solutions.⁴ The physicist, you may say, is a "Johnny-come-lately" in this respect. He is trying to impress chemists with the comparatively new subject of semiconductors, and using concepts which he has had to develop for himself for the first time; although, if he had taken the trouble to consult the chemists, he would have been told that the necessary formalism has been available for many years. One example is in the use of electrochemical potentials. Electrochemists define an electrochemical potential (denoted by μ_n) for each ion present in a solution under equilibrium conditions.⁵ In the case of the ionic equilibrium $XY \rightleftharpoons X^+ + Y^-$, the *sum* of the electrochemical potentials for X^+ and Y^- is a constant, provided that the activity of XY is fixed. In a semiconductor, physicists define a Fermi-level for holes and for electrons, and under equilibrium conditions the *difference* between the two is a constant, and is, in fact, taken to be zero. Both the electrochemical potentials and the Fermi-levels are related logarithmically to the concentration of the specified charge carrier and the formal similarity is complete except for a change of sign. Physicists go further than this, and extend the definition of Fermi-level to cover departures from thermal equilibrium,² using the same logarithmic form. In both cases, the equilibrium concentration of one carrier is described by quoting the electrochemical potential for that carrier, so that, knowing the equilibrium constant, the concentration of the other is determinate and the state of the system is completely defined. Here, the physicist deviates. Instead of quoting the analogue of the pH of the solution, he prefers to describe the equilibrium configuration by giving the ratio p/n , which is the more significant quantity when the temperature is different from the standard temperature at which pH is defined. The Fermi-level or potential (*i.e.* electrochemical potential when the system is in equilibrium) is given by

$$(\varphi_F - \varphi_{FI}) = \frac{1}{2} \frac{kT}{e} \ln (p/n) \quad (2)$$

where φ_{FI} is the value of the Fermi potential φ_F when the germanium is pure or intrinsic and $p = n$.

One final point of comparison between semiconductors and strong electrolytes may be made. For a strong electrolyte, it is permissible to ignore interionic interactions so long as the concentration is of the order of 10^{-4} molar or less. For more concentrated solutions, one must make use of some model to take account of these interactions, such as the Debye-Hückel theory.⁶ In an analogous way, interaction becomes important in germanium if the net dissolved impurity concentration is higher than about the same figure. In what follows, we shall be concerned only with the dilute solutions. In chemical terms, the activity coefficient is unity both for holes and electrons, and concentrations replace activities in the thermodynamic relations.

Having gone into some detail to point out analogies subsisting between semiconductors and electrolytes, we are now in a position to describe the simplest possible type of semiconductor interface, the p - n junction, in terms of a chemical analogy. Imagine a system consisting of an acid HA and an alkali MOH separated by a membrane permeable only to H^+ and OH^- . In order for equilibrium to be maintained, there must be a space-charge double layer at the membrane, and a Donnan membrane potential.⁷ In the same way, there must be a space-charge double layer at a p - n junction in a semiconductor, and a potential jump (V_B) given by:

$$n_1/n_2 = p_2/p_1 = \exp(eV_B/kT) \quad (3)$$

where the subscripts refer to the concentrations of the holes and electrons on the sides 1 and 2 of the junction. The relation between the barrier potential, V_B , and the space charge double layer is given by Poisson's equation. Solution of this equation shows that the extent of the potential jump depends on how rapidly the difference $N_D - N_A$ between the two impurity concentrations changes across the boundary. In germanium, the breadth of the space-charge region is of the order of 10^{-4} to 10^{-6} cm. Notice the difference in this respect between a metal and a semiconductor. In a metal, the space-charge region does not exceed a few atomic diameters in width.

The appearance of a space-charge double layer and the associated potential jump is a common feature of all surfaces. The semiconductor surface is particularly simple to deal with, because the space-charge region is so wide that one can, for many purposes, treat the region as a classical continuum in which variations in potential over distances of the order of atomic dimensions can be ignored. This case is even simpler than the Donnan equilibrium, because the solute ions are frozen into the lattice instead of being free to move. The potential jump across a p - n junction in equilibrium, however, cannot be measured directly on an external voltmeter, as can a Donnan potential difference. The physicist is not so fortunate as the chemist, who can make use of an electrode for which the half-cell potential depends on the concentration of one carrier (such as H^+), no matter whether it is the majority or the minority carrier. In a semiconductor, it turns out (see later) that the best one can do is to measure a "floating potential." This potential depends on how the minority carrier concentration differs from the equilibrium concentration and is zero when this difference is zero.

When an external potential difference is applied to a p - n junction, current flows across the interface. The applied voltage V_A adds to or subtracts from V_B , thus upsetting the equilibrium balance at the interface. If the voltages add, holes tend to drain out of the n -type region and electrons out of the p -type region. Since, in each case, these are the minority carriers they are limited in supply by the rate of thermal generation on each side. In particular, if the combined voltage $V_B + V_A$ is large enough to reduce to zero the tendencies of majority carriers to diffuse across the interface, the concentration of both carriers at the interface will approach zero, and the current saturates. This phenomenon is

analogous to that of *limiting current density*⁸ in cathodic overvoltage at an electrode on which a metal ion M^+ is being deposited, and in which ionic diffusion is the limiting process. The current across a p-n junction is of the form:

$$I = (I_{ps} + I_{ns}) [\exp (eV_a/kT) - 1] \quad (4)$$

where I_{ps} and I_{ns} are the saturation minority carrier current densities. For V_a positive, the forward direction of the rectifier, the resistance offered by the junction decreases steadily without limit until, at sufficiently large forward bias, the current is limited by ohmic (majority carrier) drops across the body of the germanium. The two saturation currents naturally depend on the lifetime of minority carriers on the two sides of the junction.

As mentioned earlier, extra minority carriers over and above the equilibrium densities may be created by shining light on to germanium. If a germanium crystal containing a p-n junction is illuminated in a region near the junction, there will be an additional supply of minority carriers which, if they reach the junction before recombining, will add to the observed current passed at a particular voltage. If L is the added current originating in this way, EQUATION 4 becomes:

$$I = I_s \left[\exp \left(\frac{eV_a}{kT} \right) - 1 \right] - L \quad (5)$$

As a particular illustration of the consequences of this equation, it will be seen that an illuminated p-n junction across which no current is flowing will be found to show a "floating potential", logarithmically related to the intensity of the incident light.² The voltage jump across the junction itself is the sum of this voltage and the internal double-layer potential difference, V_B , given by EQUATION 3. The measurement of floating potential is a way of determining how the minority carrier concentrations differs from the equilibrium case. Experimentally, it is usually easier to make measurements on an ac basis, by determining the voltage response when the p-n junction is illuminated by chopped light, the current through the junction being held constant. It may be shown that:

$$\left(\frac{\partial V_a}{\partial L} \right)_I = \frac{kT}{e} \frac{1}{I_s + I + L} = \frac{kT}{e} \frac{1}{I_s} \left(\frac{p}{p_1} \right)_n \quad (6)$$

where $(p/p_1)_n$ is the ratio of the equilibrium hole density to the actual hole density just beyond the space-charge region in the n-type material. It is also the ratio of the equilibrium and actual electron densities just inside the p region.

Now let us consider what happens to the properties of this p-n junction as the p side is made steadily less p by adding donor impurities to it, the n side being left as it is. By the time enough donor impurity has been added to change the conductivity type of the p side, and to make its conductivity equal to that of the n side, all rectifying properties will have disappeared. There will no longer be any photo-voltage across the junction, since the two sides of the junction have become identical. Now continue the process, and add more donor impurity to what was the p side until it becomes much more strongly n than the original n

side. Denote this condition by n^+ . It may now be shown that the photo-voltage is given by:

$$\left(\frac{\partial V_a}{\partial L}\right)_I = -\frac{kT}{q} \left(\frac{p}{n}\right)_n \frac{1}{I_s} \left(\frac{n}{n_1}\right)_n \quad (7)$$

One sees that this voltage is of opposite sign to EQUATION 6 and is of the order of n/p times smaller. The behaviour of the photo-voltage between the extremes represented by EQUATIONS 6 and 7 is such that, as the variable side is changed from p to n^+ , this voltage decreases monotonically from the large positive extreme EQUATION 6 to the small negative extreme EQUATION 7.

Considerable success has been obtained in the understanding of the interface between a single crystal of germanium and a gas ambient in terms of these considerations.⁹ Here, no net current crosses the surface. Changes in surface potential have been determined by measuring changes in the contact potential between germanium and a platinum reference electrode. It was found that the contact potential could be cycled between two extremes. One extreme was obtained with active oxygen, and the other with vapors containing an OH radical, such as water, carried by nitrogen. Dry nitrogen itself was neutral in its effect on contact potential. It was also found that illumination of the surface of the germanium at any point in the cycle would change the contact potential. For n -type germanium, this change was large and positive when the contact potential was near the active oxygen extreme. As the surface was taken through the cycle to the other extreme, the light effect decreased to zero and changed sign, ending up with a small negative value at the other extreme. The magnitude of the surface photo-effect at the two extremes was found to be of the form of EQUATIONS 6 and 7, where I_s is a constant, the interpretation of which we shall discuss below. For p -type germanium the effects were inverted. It will be seen that the surface photo-effect varied in exactly the same way as would the photovoltage of a junction between regions of different impurity content, as one region (corresponding to the body) is held fixed in conductivity, and the other (corresponding to the surface) is allowed to vary the whole way from p^+ to n^+ . Qualitatively, one can understand this variation in the following terms: at the one extreme (that of active oxygen), there is a layer of negative ions on the germanium surface, which must be balanced by an equal and opposite positive charge in the surface of the germanium. This means that there will be a tendency for the surface to be p^+ , whether the body of the semiconductor is n or p . At the other extreme, the ionic layer is evidently positive in sign, so that the surface will tend to be n^+ , again independently of the conductivity type of the interior. The surface photo-effect may then be understood just as for the type of junction we considered above, as follows: when light shines on the surface, the equilibrium distribution of minority carriers is upset. Since there is not time for an adjustment in ionic charge on the surface, the surface potential has to change in such a way that the total surface charge in the germanium is not altered. This is the way the photovoltage arises at a junction. The fact that the surface photoeffect at the two extremes agree with EQUATIONS 5 and 6 is quantitative evidence for the validity of this picture.

When one comes to a quantitative interpretation of the surface photoeffect in the intermediate region, there are difficulties. The trouble is that the contact potential also includes a contribution from the potential change across the adsorbed ion layer, and this potential difference may be of the same order of magnitude as the potential difference across the germanium space-charge region. In order to explain the results a model of surface states was developed. According to this model, the surface charge which compensates that of the ion layer occurs partly in donor- and acceptor-like traps on the surface (surface states). This analysis leads to the same limiting values of surface photoeffect at the extremes of contact potential, and fits the experimental observations in better with only one arbitrary parameter.

What interpretation is to be given to I_s , which has the dimensions and form of a saturation current? The analogous quantity in the theory of a p - n junction is the sum of the maximum rates at which minority carriers can be generated thermally and diffuse to the junction. In the surface case, part of I_s is the corresponding quantity for minority carriers from the body of the semiconductor. This body generation is of course closely related to the lifetime of minority carriers in the semiconductor. But minority carriers can also recombine at the surface. In chemical terms, the surface catalyses the reaction. It is known that this reaction is 'bimolecular' and it turns out that there is a contribution to I_s arising from the reverse reaction (surface generation). The total saturation current then is the sum of two such currents. The saturation current attributable to thermal generation in the body is the current that would flow by diffusion across a boundary on one side of which the minority carrier concentration is zero and on the other side of which the concentration has the equilibrium value. This current can be expressed by the product of the concentration difference and the diffusion velocity (v_d). The other saturation current can likewise be expressed in terms of a recombination velocity (v_s). It follows then that:

$$I_s = e(v_d + v_s)x \quad (8)$$

where, in this case, x is the equilibrium minority carrier density n or p whichever is appropriate. It was found experimentally that v_s did not depend appreciably on contact potential. The mechanism of surface recombination can also be understood in terms of the surface states and, from the constancy of v_s , one can draw certain conclusions regarding the distribution in energy of these surface states, which leads to a considerable simplification of the theory.

Early work with a silicon electrode in an electrolyte indicated that the potential across the space-charge layer can also be changed by applying a potential between the electrolyte and the silicon.¹⁰ A surface photoeffect is again observed. For n -type silicon, the photovoltage is large and positive when the silicon surface is biased anodically (negative surface charge-positive space charge) and becomes very small as the bias is decreased and reversed. Exactly the opposite occurs for p -type silicon. In neither, however, has the photoeffect been observed to go through zero and change sign as in the case for germanium.

Work is now in progress on silicon in a gas ambient and on the properties of

germanium in electrolytes, and it is not unreasonable to hope that in both these cases the surface properties can be understood in terms of the solid junction between regions of different conductivity, as outlined here. Also one may hope that such work may contribute to a more general understanding of potential jumps and space-charge layers at interphase boundaries.

References

1. BECKER, J. A. 1953. Life history of adsorbed atoms, ions, and molecules. *Ann. N. Y. Acad. Sci.* **58**(6):
2. SHOCKLEY, W. 1950. *Electrons and Holes in Semiconductors*. Van Nostrand. New York, N. Y.
- SHOCKLEY, W., M. SPARKS & G. K. TEAL. 1951. *p-n junction transistors*. *Phys. Rev.* **83**: 151.
3. PEARSON, G. L. & J. BARDEEN. 1948. Electrical properties of pure silicon and silicon alloys. *Phys. Rev.* **75**(879): 865.
4. REISS, H. J. 1953. Chemical effects due to the ionization of impurities in semiconductors. *J. Chem. Phys.* **21**: 1209.
5. GLASSTONE, S. 1942. *Introduction to Electrochemistry*. Van Nostrand. New York, N. Y. P. 137.
6. GLASSTONE, S. 1942. *Introduction to Electrochemistry*. Van Nostrand. New York, N. Y. P. 140.
7. DANIELS, F. 1948. *Outlines of Physical Chemistry*. Wiley. New York, N. Y. P. 556.
8. GLASSTONE, S. 1942. *Introduction to Electrochemistry*. Van Nostrand. New York, N. Y. P. 448.
9. BRATTAIN, W. H. & J. BARDEEN. 1953. Surface properties of germanium. *Bell System Tech. J.* **32**: 1.
10. BRATTAIN, W. H. 1951. *Semi-Conducting Materials*. Butterworths. London, England. P. 37.

STABILIZATION OF METAL CARBIDES BY NONMETALLIC ELEMENTS

By Harry H. Podgurski

Engineering Research Laboratory, Engineering Department, E. I. du Pont de Nemours & Company, Wilmington, Del.

In recent years, considerable data have accumulated in the literature on the kinetics of decomposition of iron, cobalt, and nickel carbides in the temperature range at which these carbides are thermodynamically unstable, relative to the metal and graphite. Because of interest in the role that these carbides play in catalysts, many of these investigations were conducted on high surface area powders.

Following the investigations of Michel and Drain¹ and of Podgurski, Kummer, Dewitt, and Emmett,² it has been established that the presence of either sulfur or oxygen in iron-cementite composites will inhibit the decomposition rate of cementite (Fe_3C). Michel and Drain treated cementite with H_2S and found that this carbide could be heated well above 600°C . with very little decomposition. At 600°C ., the untreated cementite decomposed very rapidly. According to Michel and Drain, sulfur stabilized cementite by entering the carbide lattice. Their conclusion was based upon the result that the Curie temperature of the cementite was altered by the H_2S treatment. Their pure cementite was prepared by a CO carburization of iron oxide. Podgurski, Kummer, Dewitt, and Emmett showed, on the other hand, that CO-carburized iron powders are not oxygen-free and that cementite prepared by the following reactions:



can be heated for many hours at 500°C . without incurring excessive decomposition. At 600°C ., this cementite, which contains some oxygen, does decompose rapidly. If butane or propane is substituted for carbon monoxide in the above equation, then the cementite formed during a short thermal treatment of the Fe_2C -Fe composite at 450 to 500°C . will react readily on continued thermal treatment at 500°C . to form free iron and carbon.

This paper proposes that both sulfur and oxygen may readily inhibit the decomposition of iron, cobalt, and nickel carbides by blocking the nucleation of graphite at the surface of these powder composites: in other words, that the nonmetallic elements, when chemically combined with the surface, block the permeation of carbon through the surface on which the decomposition product, graphite, is nucleated.

This proposal is substantiated by both thermodynamic considerations and further experimental work on the effect of sulfur on the thermal stability of Co_2C .

Experimental

The objective sought in the preparation of the samples for this investigation was to synthesize a composite sample of cobalt and cobalt carbide (Co_2C) whose carbidic carbon content could be established and whose external surface area could be measured by gas adsorption.

The samples were contained in Pyrex glass vessels equipped with stopcocks and a bank of cassettes for removing X-ray samples under vacuum. Reduction, carburization, chemisorption studies, controlled sulfiding, and surface area measurements were made *in situ*. As a starting material, a sample of cobalt oxide containing approximately 0.2 per cent silica was reduced with dry hydrogen at temperatures (stepwise) from 200 to 405° C. over a period of three days. This sample was then treated with carbon monoxide containing 1.5 per cent hydrogen, observing the schedule:

Carburizing temperature	Total time at temp.	Per cent wt. increase
	<i>hr.</i>	
195° C.	17	0.12
225-230° C.	62.5	1.5
225-230° C.	112.5	2.5
225-230° C.	160.5	3.7
225-230° C.	274.5	5.3*

* Three samples of this material, prior to any further treatment, are identified in TABLE I as Co (5.3 per cent C)-1, 2, and 3.

The extent of carburization of the cobalt was followed gravimetrically. After the last carburizing treatment, the amount of carbon incorporated into the sample as carbidic carbon (Co_2C), as distinguished from the free carbon, was established by a low temperature reduction (2-hour reduction at 300° C.). Elemental carbon or graphite is not reduced or converted to methane readily at 300° C.

The carbide prepared was identified by X-ray analysis, as were the subsequent phase transformations that followed the vacuum-heat treatment. Various heat treatments were conducted at 348, 400, and 500° C. at pressures less than 10^{-5} mm. of mercury. During these heat treatments, the samples were continuously pumped to remove gaseous reaction products. X-ray diffraction patterns were taken of the reduced cobalt and the carburized cobalt before and after vacuum-heat treatment.

Surface area measurements (B.E.T. method⁶) employing nitrogen as the adsorbate at -195° C. were also conducted on the reduced metal powder and on the carburized powder sample before and after thermal decomposition studies. To determine whether the small trace of silica or other contaminants in this sample covered an appreciable fraction of the surface, CO chemisorption studies were conducted on the reduced sample. This method of characterizing iron and cobalt surfaces has been described^{4, 5, 6} on several occasions.

In the preliminary experiments, the sulfided carbide-containing cobalt samples were prepared by simply maintaining an atmosphere of H_2S over the sample at a particular temperature, and cooling to room temperature in the

same atmosphere. The sample was then removed and made ready for X-ray analysis or further vacuum-heat treatment.

In the more critical experiments, a sample was exposed to a measured quantity of H_2S . The H_2S was pumped through the sample in a vacuum-tight glass system with a magnetically operated glass pump as the sample temperature was gradually raised from -78°C . to 200°C . The amount of unreacted H_2S , at any given time, was determined by freezing it in a refrigerated trap (-195°C .) while the uncondensed hydrogen, produced by the reacted H_2S , was pumped from the system. The unreacted H_2S was then volatilized and measured volumetrically. This procedure effected uniform distribution of the sulfide over the surface of the sample. In this experiment, most of the H_2S had reacted before the sample temperature reached 100°C . The volumetrically measured quantity of reacted H_2S was finally confirmed gravimetrically.

Results

The X-ray diffraction pattern obtained for the cobalt carbide prepared by CO carburization of reduced cobalt powder was identical to the carbide whose structure was established by Drain and Michel.⁷ Within the limits of experimental error, it appears that neither the sulfiding treatments nor subsequent heat treatments altered the structure of the retained or undecomposed cobalt carbide.

At temperatures as high as 500°C . (2-hour vacuum-heat treatment), the stabilizing influence of sulfur on Co_2C could be detected, *i.e.*, some Co_2C had not decomposed. At 400°C . (2-hour vacuum-heat treatment), the effect was pronounced. Increasing the sulfur content of the composites to the point where cobalt sulfide could be detected by X-ray analysis did not appear to inhibit the carbide decomposition reaction any more than when an equivalent of one to three monolayers of cobalt sulfide had been formed on the sample. Judging from the relative intensity* of the cobalt and cobalt carbide lines in the X-ray diffraction patterns obtained for the sulfided samples [Co (5.3 per cent C) — 2S and 2SH] it appeared that very little decomposition had occurred during the 7-hour vacuum-heat treatment (REACTION 3, TABLE 1†). Compare FIGURE 1 a, b, and c. On the other hand, observed area changes revealed that considerable decomposition must have occurred. This observation suggests that some of the tiny crystallites of undecomposed Co_2C , which give very poor diffraction patterns at low concentrations, had grown at the expense of others. At 348°C ., the decomposition reaction, which will go to completion when no sulfur is present, is apparently slowed down sufficiently so that an observable crystal growth was probably observed. This, of course, implies that sulfur does not affect the mobility of carbon across the cobalt-cobalt carbide interface.

The stabilizing influence of sulfur on the thermal stability of Co_2C was unambiguously established by X-ray diffraction analyses. These analyses, however, gave no indication as to how stabilization was effected. Furthermore, they gave no evidence of the nature or distribution of the graphite formed from the

* By visual examination.

† REACTIONS 1 through 8 are shown in TABLE 1.

decomposition of the Co_2C . The following surface studies are believed to give such information.

The starting samples of reduced cobalt powder prepared by the reduction of cobalt oxide with hydrogen will yield areas, as measured by gas adsorption, between 3 and 4 sq.m./g. of Co. The variations in the initial area values will depend upon the reduction procedure (hydrogen space velocity through the powder, schedule, and temperature). The samples of cobalt employed in a good portion of this work had a measured area of 3.3 sq. m./g. of Co.

In an attempt to characterize further the surface of our original reduced cobalt surface, a series of CO adsorption studies was conducted at -195°C . on the reduced sample. The ratio of the volume of CO (V_{CO}) chemisorbed at -195°C . to the nitrogen monolayer (V_m) amounted to 0.62 (V_{CO}/V_m). A

TABLE 1
SURFACE AREA CHANGES FOR REACTIONS INVOLVING Co- Co_2C COMPOSITES

Reaction no.	Reactant	Treatment	Product
1	<i>Co (Reduced)</i> [A*: 3.3 sq.m./g. of Co $V_{\text{CO}}/V_m = 0.63$]	Carburized with CO ($1\frac{1}{2}\%$ H_2) 195 to 230°C .	<i>Co (5.3% C)—2 & 3</i> [3.74% C as Co_2C 1.56% C as graphite A: 8.8 ± 0.2 sq.m./g. of Co Co as Co_2C , 40.6%; Co, 59.4%]
2	<i>Co (5.3% C)—2</i> [A: 8.6 sq.m./g. of Co]	2.16 cc. (S.T.P.) of H_2S reacted per g. of Co -78° to 200° C.	<i>Co (5.3% C)—2S</i> [Co as Co_2C : 40.0 to 40.6% Co as Co: 58.8 to 59.4% Co as Co_9S_8 , 0.6% A: 7.9 sq.m./g. of Co]
3	<i>Co (5.3% C)—2S</i> [A: 7.9 sq.m./g. of Co]	Vacuum heat- treatment 7 hr. at 348°C .	<i>Co (5.3% C)—2SH</i> [A: 10.2 sq.m./g. of Co X-ray analysis showed Co_2C still present]
4	<i>Co (5.3% C)—3</i> [A: 9.0 sq.m./g. of Co]	Vacuum heat- treatment 7 hr. at 338° to 348°C .	<i>Co (5.3% C)—3H</i> [A: 11.9 sq.m./g. of Co X-ray analysis showed no Co_2C present]
5	<i>Co (5.3% C)—3H</i> [A: 11.9 sq.m./g. of Co]	Vacuum heat- treatment 15 hr at 390°C .	<i>Co (5.3% C)—3HH</i> [A: 12.03 sq.m./g. of Co]
6	<i>Co (5.3% C)—3HH</i> [A: 12.03 sq.m./g. of Co]	H_2 reduction 2 hr at 300°C .	<i>Co (5.3% C)—3HHR</i> [No. wt. loss observed A: 12.08 sq.m./g. of Co]
7	<i>Co (5.3% C)—3HHR</i> [A: 12.08 sq.m./g. of Co]	H_2 reduction 15 hr. at 396°C .	<i>Co (0.0% C)</i> [A: 3.3 sq.m./g. of Co]
8	<i>Co (5.3% C)—2SH</i> [Co_2C present]	H_2 reduction 2 hr. at 300°C .	<i>Co (5.3% C)—2SHR</i> [Co_2C still present (by x-ray analysis)]

* Arca.

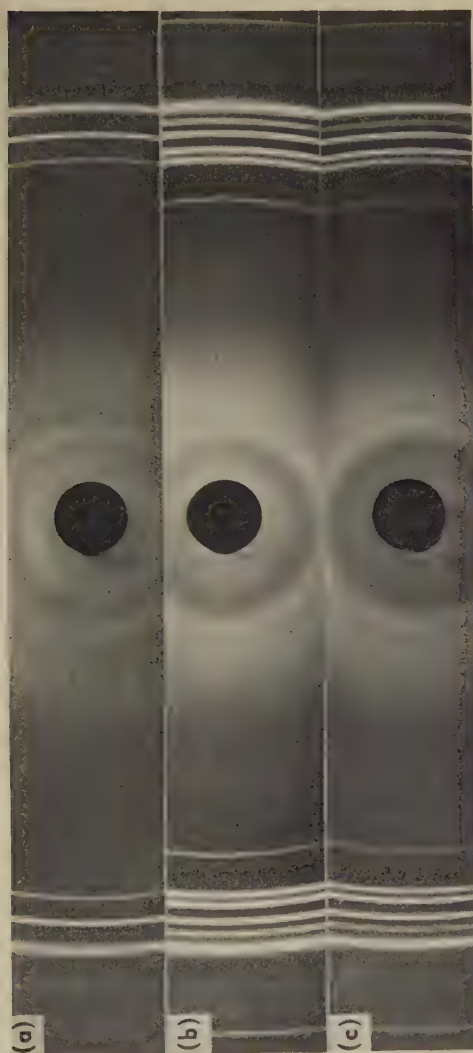


FIGURE 1. N-ray powder patterns (filtered Co radiation). (a) Co (5.3 per cent C)—3H; (b) Co (5.3 per cent C)—2S; (c) Co (5.3 per cent C)—2 SH.

value of 0.65 was observed by Anderson, Hall, and Hofer,⁸ on a relatively pure sample of cobalt metal powder. As these values are quite similar, we assumed that the surface of our own sample was not badly contaminated.

Following the complete carburization schedule reported above, a surprisingly high value of 8.8 ± 0.2 sq.m./g. of Co was obtained (see TABLE 1, REACTION 1). Comparing this value with the one obtained on the reduced cobalt sample, it was concluded that a considerable amount of finely dispersed carbon or graphite had deposited on the sample surface during the carburization treatment. Carburization of cobalt and iron powders does not effect a noticeable change in area if no graphite or elemental carbon is formed.

Having assumed from our area measurements that the 5.3 per cent C incorporated into the sample by carburization was in part graphitic carbon, a low-temperature hydrogen reduction, which readily removes carbidic carbon (Co_2C) but not graphitic carbon, was conducted on part of the initial sample to establish this point. A 2-hour reduction with dry hydrogen removed only 3.74 per cent of the 5.3 per cent carbon, which is consistent with our assumption that the high area obtained on the carburized sample was owing to the presence of graphitic carbon. This criterion is supported by the following results. REACTIONS 5 and 6, TABLE 1, were accompanied by no significant change in area or weight, indicating the absence of any carbidic carbon and the low reactivity of the graphitic carbon at 300°C ., whereas in the course of REACTION 4 the area changed and, in the case of REACTION 7, both values changed. During REACTION 4 elemental carbon was formed, and in REACTION 7 the high area attributed to elemental carbon disappeared as it was removed at 396°C . by reduction with hydrogen.

Turning our attention to REACTION 2 we shall now consider the area change effected by the controlled sulfiding treatment. The drop in area on treating the carburized sample with H_2S yields information regarding the distribution of sulfide ion over the sample. Had all the H_2S reacted completely with all of the cobalt in only a few particles of the powder, the drop in area would have amounted only to 0.02 sq.m./g. of Co as compared to the observed value 0.7 sq.m. g. of Co. This calculated drop (0.02 sq.m./g. of Co) in area assumes that the reacted particles (0.6 per cent of the cobalt in the sample) would no longer contribute to the measurable area. The area of the reduced sample was chosen as the basis for calculating the amount of H_2S to be added to a sample to produce, if completely reacted, an equivalent of one to two monolayers of cobalt sulfide on the surface of the carburized cobalt. The surface created by the nucleation of graphite was not included as it was assumed to be unreactive. A value of 3.3 sq.m./g. of Co was assumed for the effective or reactive carburized cobalt surface. This value could be in error by the amount of surface not accessible at the interface area between the nucleated graphite and the carburized cobalt. Then 2.16 cc. of H_2S /g. of Co was reacted with this sample. The resulting coverage was estimated to be between one and three monolayers. Obviously, the calculated coverage will also depend upon the packing factor assumed for the sulfide ion ($\text{S}^{=}$).

Probably the most interesting result arising from the surface studies was the

significant change in area observed during course of REACTIONS 3 and 4.* As this magnitude of area change is not always observed during thermal decomposition of carburized Co and Fe powders, and as all the factors which control the nature of the carbon deposit have not been established, we shall not dwell on the reasons for this effect except to say that the nucleation sites of the graphite deposit were established as occurring on the external surface of the composites. That the change in area can be attributed to the nucleation and growth of a high surface area deposit of graphite on the surface was further confirmed with the electron microscope. These results will be considered later.

Now the heat treatment effected in REACTION 3, as mentioned previously, did not cause the disappearance of the strong Co_2C lines in the X-ray diffraction pattern. In this case, a 29 per cent increase in area was observed. Comparing this value with the 32 per cent increase observed during the course of REACTION 4 for the unsulfided sample in which no Co_2C could be detected, we find it necessary to assume that only a very small fraction of the Co_2C present had decomposed and that the resulting graphite formed was more dispersed than that generated on the untreated composite. An alternative explanation that is equally reasonable and was previously mentioned supposes that, although a considerable amount of carbide had decomposed, the decomposition reaction had been slowed down sufficiently to allow some observable crystal growth of the Co_2C phase.

Electron microscope studies were undertaken to confirm our original assumption that graphitization should occur preferentially on the outer surface of our carburized powders. Our calculated values (based on area measurements, B.E.T. method) for the average particle size of the graphite phase formed during heat treatment and during carburization led us to believe that the resolving power of the electron microscope should make it possible to reveal these particles on the periphery of the particles of carburized cobalt. The electron micrographs shown in FIGURES 2 a, b, and c confirmed our original assumption. The hairlike deposit on the surface of these composites was identified by electron diffraction studies as graphitic carbon. No cobalt was detected in this surface deposit.

Discussion

To get at the working mechanism involved in the process by which metastable metal carbides decompose, we shall describe from the viewpoint of thermodynamics the reactions that appear to be involved during the induction periods that have been observed^{16, 17} to precede the carbide decomposition reactions.

In TABLE 2 we have listed the standard free energies of formation for Co_2C , Fe_2C (Hägg), Fe_3C , and Ni_3C at various temperatures. It is apparent from the data that all of these carbides are unstable relative to the elements, at least up to 700° C. and, on the basis of these values for the standard free energy of formation, they show remarkable similarity as contrasted to the observed stability based on decomposition rate data. Apparently, if we are to correlate any thermodynamic properties with the observed stability, we must examine our system more carefully.

* We have recently obtained the same results on silica-free (less than 0.02 per cent SiO_2) cobalt powder.

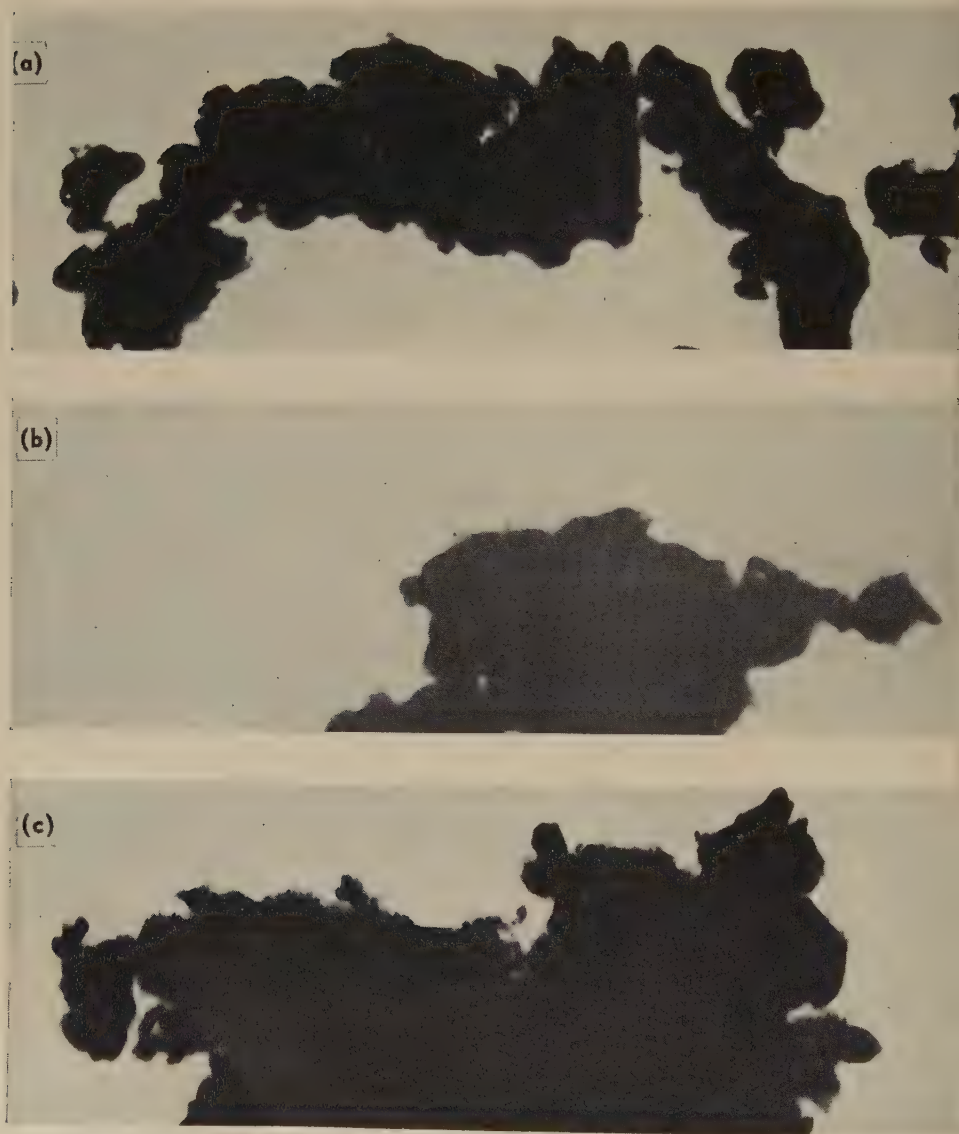


FIGURE 2. Electron micrographs (50,000 $\times$). (a) (0.0 per cent C); (b) Co (5.3 per cent C)—2S; (c) (5.3 per cent C)—2SH.

In the following treatment, it is assumed that the sulfur or oxygen functions by blocking the nucleation of graphite at the surface of these composites, and that the energies of chemical binding of these elements on the surface follow the same relative order as the corresponding bond energies in the respective bulk metal oxides and sulfides. It is quite apparent that if, during the course of vacuum-heat treatment, particularly during the observed induction period, the nonmetallic elements can be removed by some side reaction, then the nucleation of graphite might be accelerated. The feasibility of various side reactions will now be considered from a thermodynamic point of view. Presumably, these assumptions will bear some merit, providing a correlation follows.

The first set of side reactions considered was that which would yield elemental oxygen or sulfur. In general, it can be said that removal of oxygen is less likely than the removal of sulfur by the straight thermal decomposition of the respective oxide or sulfide of nickel, cobalt, or iron. These side reactions seem unreasonable, for they would predict that oxygen is more stabilizing than sulfur. This is contrary to our observations. Furthermore, the calculated values for the equilibrium partial pressures of oxygen and sulfur resulting from thermal decomposition of the respective metal oxides and sulfides in the temperature range 500 to 1000° K. are so low that these reactions would appear to be very unlikely prospects.

It remains to explore any other reaction processes that will effectively remove oxygen or sulfur as a gaseous product from the composites during vacuum heat-treatment. A few such reactions are listed below in general form:

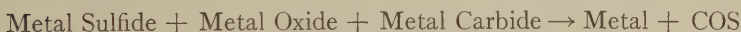


TABLE 2

FREE ENERGY OF FORMATION OF IRON, NICKEL, AND COBALT CARBIDES, ΔG°_f

$3\text{Fe} + \text{C} (\beta \text{ graphite}) \rightarrow \text{Fe}_3\text{C}^8;$		$2\text{Fe} + \text{C} (\beta \text{ graphite}) \rightarrow \text{Fe}_2\text{C} (\text{Hägg})^9$	
T°K.	ΔG°_f (cal./mole)	T°K.	ΔG°_f (cal./mole)
298	+4,441	450	+3,757
400	4,000	500	3,613
500	3,447	550	3,570
600	2,850	600	3,360
700	2,250	650	3,230
800	1,680		
900	0*		
1000	0*		
$3\text{Ni} + \text{C} (\beta \text{ graphite}) \rightarrow \text{Ni}_3\text{C}^{10};$		$2\text{Co} + \text{C} (\beta \text{ graphite}) \rightarrow \text{Co}_2\text{C}^{10}$	
500	+7,127	500	+2,400
600	+4,653	600	+1,800*
1000	0*	1000	0*

* Estimated values assumed for the purpose of calculating the ΔG°_f values in TABLE 3.

TABLE 3*

EQUILIBRIUM PARTIAL PRESSURES OF CO, COS, AND CS₂ CALCULATED FOR REACTIONS INVOLVING COMPOSITES OF METAL OXIDES, SULFIDES, AND CARBIDES

$\text{Fe}_3\text{C} + \frac{1}{4} \text{Fe}_3\text{O}_4 = \text{CO} + 3\frac{3}{4} \text{Fe}$		$2\text{FeS} + \text{Fe}_3\text{C} = 5\text{Fe} + \text{CS}_2$	$\text{Fe}_3 + \frac{1}{4} \text{Fe}_3\text{O}_4 + \text{FeS} = \text{COS} + 4\frac{3}{4} \text{Fe}$
<i>T</i> °K.	<i>P</i> _{atm.} (CO)	<i>P</i> _{atm.} (CS ₂)	<i>P</i> _{atm.} (COS)
500	1.3×10^{-7}	$<10^{-10}$	1.2×10^{-14}
600	3.1×10^{-5}	$<10^{-10}$	2.5×10^{-11}
700	1.7×10^{-3}	$<10^{-10}$	6.2×10^{-9}
800	2.8×10^{-2}	$<10^{-10}$	3.5×10^{-7}
900	1.5×10^{-1}	$<10^{-10}$	4.8×10^{-6}
1000	1.3	7.9×10^{-10}	8.4×10^{-5}
$\text{Ni}_3\text{C} + \text{NiO} = \text{CO} + 4\text{Ni}$		$\text{Ni}_3\text{S}_2 + \text{Ni}_3\text{C} = 6\text{Ni} + \text{CS}_2$	$\text{Ni}_3\text{C} + \frac{1}{2} \text{Ni}_3\text{S}_2 + \text{NiO} = 5\frac{1}{2} \text{Ni} + \text{COS}$
500	9.2×10^{-2}	1.5×10^{-10}	1.1×10^{-7}
600	0.7	1.9×10^{-15}	2.3×10^{-5}
1000	$6.1 \times 10^{+2}$		
$\text{CoO} + \text{Co}_2\text{C} \rightarrow 3\text{Co} + \text{CO}$		$\frac{3}{4} \text{Co}_9\text{S}_8 + \text{Co}_2\text{C} \rightarrow 4\frac{1}{4} \text{Co} + \text{CS}_2$	$\text{CoO} + \text{Co}_2\text{C} + \frac{1}{8} \text{Co}_9\text{S}_8 = \text{COS} + 4\frac{7}{8} \text{Co}$
500	4.2×10^{-4}	7.2×10^{-26}	4×10^{-11}
600	2.3×10^{-2}	6.2×10^{-20}	3.3×10^{-8}
700	2.9×10^{-1}	—	4.2×10^{-6}
800	2.2	—	1.1×10^{-4}
900	17.3	—	2.8×10^{-3}
1000	88.6	2.8×10^{-8}	3.6×10^{-2}

* Calculations based on data taken from Refs. 9, 10, 12, 13, 14, and 15.

In TABLE 3 we have compiled our calculated values of partial pressures of gaseous reaction products for a series of specific reactions involving Fe, Co, and Ni. The accuracy of these calculated values was definitely limited by the accuracy of available data on the free energy of formation of the metal sulfides and to some extent of the carbides. As our interpretation depends upon trends only, a factor of 10 in partial pressure values should not be considered significant. It will be assumed that, for a given temperature, the most rapid contaminant-removal reaction will be that reaction which shows the highest equilibrium partial pressure of reaction product containing combined oxygen or sulfur.

On this basis, we can make the general statement that sulfide-treated carbides should be more stable than the same carbide containing oxygen as the contaminant. The evidence taken from the literature on the iron system is in agreement with this observation. If our assumptions are correct, then it follows that sulfide-treated cobalt and nickel carbides should show a greater resistance to thermal decomposition than the respective untreated CO-carburized samples. As no data were available on the cobalt or nickel system, we chose to test the assumption on the cobalt system.

To extend the test of our proposed assumptions, we might attempt to compare the properties of all three metal carbides prepared by the CO carburization

method. The data in TABLE 3 suggest that iron carbide should resist thermal decomposition better than cobalt or nickel carbide when prepared in the same manner, whereas cobalt carbide should behave very similar to nickel carbide. Again the literature substantiates our assumption, even though we are undoubtedly stretching our original intentions, for the reason that it is difficult to obtain even comparable surface conditions (surface area, *etc.*). To attempt to compare the three sulfided CO-carburized metals might be misleading for the same reasons.

It should be kept in mind that the activation energy for the decomposition reaction of an oxygen- or sulfide-free carbide is not assumed to be zero and most probably will also depend on whether one is dealing with only the carbide phase or a metal-metal carbide composite.

Summary

The evidence presented does not establish the location of the sulfur and oxygen which lowers the decomposition rate of the carbides in these metal-metal carbide composites. In our opinion, the data favor a surface effect rather than the bulk effect suggested by Michels and Drain, as based on their interpretation of their own thermal magnetic studies. One could possibly account for the observed change in the Curie temperature of iron carbide by the strain introduced into the carbide lattice as a surface oxide or sulfide is formed during the preparation of the carbide.

Further evidence that the effect observed is truly one to be attributed to a surface factor follows from some observations recently brought to our attention by Professor J. A. Kitchener of Imperial College, University of London. In dealing with a totally different type of system, Spratt and Kitchener¹¹ made the observation that sulfur in molten iron definitely inhibits the dissolution of cementite and likewise inhibits graphitization on solidification. They stated that it was unlikely that a small amount of sulfur which is practically insoluble in cementite should so displace the thermodynamic properties of Fe-C systems as to make Fe_3C thermodynamically stable. They further pointed out the Fe_3C has not been observed to form more spontaneously and continuously at any temperature as a stable phase in the presence of sulfur. In our own studies on cobalt powders the solubility of sulfur in the carbide at the temperature under consideration must be exceedingly small.

Acknowledgment

The X-ray diffraction analyses were made by A. I. Petrovich and the electron micrographs were taken by R. H. Fisher.

References

1. MICHEL, A. & J. DRAIN. 1951. Pittsburgh Conference on X-Ray and Electron Diffraction.
2. PODGURSKI, H. H., J. T. KUMMER, T. W. DEWITT & P. H. EMMETT. 1950. J. Am. Chem. Soc. **72**: 5382-5388.
3. HÄGG, G. 1934. Z. Krist. **89**: 92.
4. EMMETT, P. H. & S. J. BRUNAUER. 1935. J. Am. Chem. Soc. **51**: 1754.
5. BRUNAUER, S. & P. H. EMMETT. 1940. J. Am. Chem. Soc. **62**: 1933.

6. EMMETT, P. H. 1945. *Ind. Eng. Chem.* **37**: 639.
7. DRAIN, J. & M. MICHEL. 1951. *Bul. Soc. Chim. France*. **18**: 517.
8. ANDERSON, R., K. HALL & L. J. HOFER. 1948. *J. Am. Chem. Soc.* **70**: 2465.
9. BROWNING, L. C., T. W. DEWITT & P. H. EMMETT. 1950. *J. Am. Chem. Soc.* **72**: 4215.
10. BROWNING, L. C. & P. H. EMMETT. 1952. *J. Am. Chem. Soc.* **74**: 1682.
11. SPRATT, D. A. & J. A. KITCHENER. 1952. *J. Iron Steel Inst. London, England*. **172**: 383-386.
12. THOMPSON, D. M. 1952. *The Total and Free Energies of Formation of the Oxide of Thirty Metals*. Electrochem. Soc. New York, N. Y.
13. RICHARDSON, F. D. & J. H. E. JEFFES. 1948. *J. Iron Steel Inst. London, England*. **160**: 261-270.
14. RICHARDSON, F. D. & J. H. E. JEFFES. 1952. *J. Iron Steel Inst. London, England*. **171**: 165-175.
15. KIRKPATRICK, W. J. 1951. Nickel sulfide catalysts. In *Advances in Catalysis*. **3**: 331. Academic Press. New York, N. Y.
- 16 & 17. HOFER, L. J. E., E. M. COHN & W. C. PEBLES. 1949. *J. Phys. & Colloid Chem.* **53**: 661, **54**: 1161.

MONOGRAPHIC PUBLICATIONS
OF
THE NEW YORK ACADEMY OF SCIENCES
(LYCEUM OF NATURAL HISTORY, 1817-1876)

(1) The ANNALS (octavo series), established in 1823, contain the scientific contributions and reports of researches, together with the records of meetings of the Academy. The articles which comprise each volume are printed separately, each in its own cover, and are distributed immediately upon publication. The price of the separate articles depends upon their length and the number of illustrations, and may be ascertained upon application to the Executive Director of the Academy.

Current numbers of the ANNALS are sent free to all Members of the Academy desiring them.

(2) The SPECIAL PUBLICATIONS established in 1939, are issued at irregular intervals as cloth-bound volumes. The price of each volume will be advertised at time of issue.

(3) The MEMOIRS (quarto series), established in 1895, are issued at irregular intervals. It is intended that each volume shall be devoted to monographs relating to some particular department of Science. Volume I, Part 1, is devoted to Astronomical Memoirs, Volume II to Zoological Memoirs. No more parts of the Memoirs have been published to date. The price is one dollar per part.

(4) The SCIENTIFIC SURVEY OF PORTO RICO AND THE VIRGIN ISLANDS (octavo series), established in 1919, gives the detailed reports of the anthropological, botanical, geological, paleontological, zoological, and meteorological surveys of these islands.

Subscriptions and inquiries concerning current and back numbers of any of the publications of the Academy should be addressed to

EXECUTIVE DIRECTOR
The New York Academy of Sciences
2 East Sixty-third Street
New York 21, N. Y.

DATE DUE